

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 24, 2004, 20:09:36 ; Search time 5876 Seconds
(without alignments)
4772.455 Million cell updates/sec

Title: US-10-086-464-2

Perfect score: 3453

Sequence: 1 MSSAPSGTGGPPSPNST.....REMEMGKIKRTGGYSGPSL 647

Scoring table: BL0SUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xip
-Q=/cgn2/1/USPTO_spool_p/US10086464/runat_23042004_083015_5613/app_query.fasta_1.839
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10086464 @CGN 1.1 5265 @runat_23042004_083015_5613 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*

- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pin.*
- 35: em_htg_rod.*
- 36: em_htg_mam.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	3453	100.0	1944	6	AX088876 Sequence
2	3453	100.0	1944	6	AX825703 Sequence
3	3453	100.0	2189	6	AX825705 Sequence
4	3453	100.0	2189	8	AY028699 Braesica
5	2936.5	85.0	1959	6	AX825738 Sequence
6	2936.5	85.0	1959	8	BT008400 Arabidops
7	2936.5	85.0	2098	8	AY128792 Arabidops
8	2936.5	85.0	2116	8	AY056788 Arabidops
9	2936.5	85.0	2188	8	AY059901 Arabidops
10	2936.5	85.0	2190	8	AY093065 Arabidops
11	2934.5	85.0	1959	8	BT008409 Arabidops
12	2934.5	85.0	2257	8	AF370509 Arabidops
13	2885	83.6	2324	8	AY089024 Arabidops
14	2622	75.9	79706	8	AB020746 Arabidops
15	1841.5	53.3	2307	8	AK103247 Oryza sat
16	1821.5	52.8	1812	6	AX825735 Sequence
17	1812	52.5	80393	8	AP000382 Arabidops
C	1748.5	50.6	1902	6	AX088882 Sequence
18	1748.5	50.6	1902	6	AX825712 Sequence
19	1748.5	50.6	1939	6	AX088881 Sequence
20	1748.5	50.6	1939	6	AX825711 Sequence
21	1748.5	50.6	2013	8	BT005955 Arabidops
22	1746	50.6	2013	8	AK118488 Arabidops
23	1746	50.6	2320	8	AX088885 Sequence
24	1687	48.9	2025	6	AX825715 Sequence
25	1687	48.9	2025	6	AX088884 Sequence
26	1687	48.9	2104	6	AX825714 Sequence
27	1687	48.9	2104	6	AX825736 Sequence
28	1668	48.3	1515	6	AX825737 Sequence
29	1637.5	47.4	1488	6	AX224526 Oryza sat
30	1574.5	45.6	1185	8	AY224526 Oryza sat
31	1567.5	45.4	2364	8	AK064397 Oryza sat
32	1567.5	45.4	2371	8	AK102832 Oryza sat
33	1507	43.6	82356	8	AC006135 Arabidops
34	1507	43.6	92861	8	AC006439 Arabidops
35	1480.5	42.9	125021	8	AC007504 Arabidops
36	1479.5	42.8	114367	8	ATF10M10 Arabidops
37	1479.5	42.8	125502	8	ATT4L20 Arabidops
38	1479.5	42.8	195921	8	ATCHRIV81 Arabidops
39	1459.5	42.3	2158	8	AY113877 Arabidops
40	1459.5	42.3	3064	8	AY035076 Arabidops
41	1456	42.2	2046	8	AY113039 Arabidops
42	1456	42.2	2616	8	AF424623 Arabidops
43	1456	42.2	2735	8	AY075681 Arabidops
C	1434.5	41.5	120465	2	AC137001 Oryza sat
44	1431	41.4	3070	8	AK110196 Oryza sat
45					

ALIGNMENTS

RESULT 1

AX088876	AX088876	1944 bp	DNA	linear	PAT 17-MAR-2001
LOCUS	Sequence 1	from Patent WO0114563.			
DEFINITION	AX088876				
ACCESSION	AX088876				
VERSION	AX088876.1	GI:13397639			

ORIGIN					
Alignment Scores:					
Pred. No.:	4.02e-76	Length:			1944
Score:	3453.00	Matches:			647
Percent Similarity:	100.00%	Conservative:			0
Best Local Similarity:	100.00%	Mismatches:			0
Query Match:	100.00%	Indels:			0
DB:	6	Gaps:			0

US-10-086-464-2 (1-647) x AX088876 (1-1944)

Qy		1	MetsSerAlaProSerProGlyThrGlySerProSerProSerProSerAsnSerThr	20
Db		1	ATGTCTCGGGCGCGTTCGCGGGACGTGTTGCCTCATCTCCACCATCAAACTCCACA	60
Qy		21	ThrThrThrProProAlaSerAlaProProThrThrProSerSerProProPro	40
Db		61	ACCACCACTCCTCTCCAGCTTCGCGTCTCTCCCACCACACTCTTCTCTCTCGCGCG	120
Qy		41	ProSerThrIleProThrSerProProProSerSerArgSerThrProSerAlaProPro	60
Db		121	CCATCCACTATTCCGACATCTCCTCCTCTTCTTGCTACACCTTCTGCTCCTCCT	180
Qy		61	ProSerProProThrProSerThrProGlySerProProProLeuProGlnProSerPro	80
Db		181	CCATCTCCACCACATCCCATCTACGCCGGGATCTCCACTCTCTCTTCCTCAGCCGTCTCCA	240
Qy		81	ProAlaProThrThrProGlySerProProAlaProValThrProProThrArgAsnPro	100
Db		241	CCCCCTCAACTACGCCGGGATCTCCACCCGCACCTGTTACTCCTCTCTCTCGAAGCCCT	300
Qy		101	ProProSerValProGlyProProProSerAsnProSerArgGluGlyGlySerProArgPro	120

Db 841 TTAGACAAGCGGGTTCGGTTACGTGCACAAAGGTGTGTTCCCTAGTGGGAAGAAGTT 900
Qy 301 AlaValLysGlnLeuLysValGlySerGlyGlnGlyGluArgGluPheGlnAlaGluVal 320
Db 901 GCTGTGAACAGAGTTGAAAGTTGGAGTGGTCAGGGAGAGAGGAGTTTCAGGCAGAGGTT 960
Qy 321 GluIleIleSerArgValHisArgHisLeuValSerLeuValGlyTyrCysIleAla 340
Db 961 GAGATCATCAGCAGAGTTCCACACAGGCATCTGGTGCTCTGTGTGGTTATTGCATGCC 1020
Qy 341 GlyAlaLysArgLeuLeuValTyrGluPheValProAsnAsnAsnLeuGluLeuHisLeu 360
Db 1021 GGTGCCAAAGATTGCTGTATGAGTTTGTCTTAACAACAATCTCGAGCTTCACCTC 1080
Qy 361 HisGlyGluGlyArgProThrMetGluTrpSerThrArgLeuLysIleAlaLeuGlySer 380
Db 1081 CATGCGAGGACCGCCTACAAATGGAATGGAGCACCAATTCGAAGATTGCTCTTGGATCT 1140
Qy 381 AlaLysGlyLeuSerTyrLeuHisGluAspCysAsnProLysIleIleHisArgAspIle 400
Db 1141 GCTAAAGGACTTCTTATCTTCATGAAGATTGCAATCTCTAAATCATTCACCGTGATATC 1200
Qy 401 LysAlaSerAsnIleLeuIleAspPheLysPheGluAlaLysValAlaAspPheGlyLeu 420
Db 1201 AAGGCTTCAACATATTGATAGATTCAAGTTTGAAGCTAAGTTGCTGATTTTGGTCTT 1260
Qy 421 AlaLysIleAlaSerAspThrAsnThrHisValSerThrArgValMetGlyThrPheGly 440
Db 1261 GCTAAGATTGCTTCTGATACAAACACGCGATGTATCAACACGCTGTGATGGGAACCTTTGGG 1320
Qy 441 TyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPheSer 460
Db 1321 TACTTGGCTCCGGAATAGCTGCAAGCGGAAGCTCACGGAGAAGTCTGACGTTTTCTCA 1380
Qy 461 PheGlyValValLeuLeuGluLeuIleThrGlyArgArgProValAspAlaAsnVal 480
Db 1381 TTTGGCGTTGTCTTTGGAGCTCATTTACTGGAGCTCGACCGTGTGATGCCAACAATGTC 1440
Qy 481 TyrValAspAspSerLeuValAspTrpAlaArgProLeuLeuAsnArgAlaSerGluGln 500
Db 1441 TATGTAGATGACAGCTTAGTGTGACTGGCGCAGCACCATTGCTTAAACCGAGCATCTGAGCAA 1500
Qy 501 GlyAspPheGluGlyLeuAlaAspAlaLysMetAsnGlyTyrAspArgGluGluMet 520
Db 1501 GGAGACTTTGAGGGTTTGGCTGATGCCAAGATGAATAATGGGTATGACAGAGAGAGATG 1560
Qy 521 AlaArgMetValAlaCysAlaAlaAlaCysValArgHisSerAlaArgArgProArg 540
Db 1561 GCTGCATGGTTGCTTGTGCTGGGCTTGTGTTCGCCATTTCAGCTCGCGCGCAGACCTGCG 1620
Qy 541 MetSerGlnIleValArgAlaLeuGluGlyAsnValSerLeuSerAspLeuAsnGluGly 560
Db 1621 ATGAGCCAGATTGTGCGTGGCTTAGAAGGAATGTATCACTGTCCAGATCTTAAACGAAGGG 1680
Qy 561 MetArgProGlyGlnSerAsnValTyrSerSerTyrGlyGlySerThrAspTyrAspSer 580
Db 1681 ATGAGCCAGGTCAAGCAATGTATACAGCTCATACGGAGGAAGCACCGATTATGACTCG 1740
Qy 581 SerGlnTyrAsnGluAspMetLysLysPheArgLysMetAlaLeuGlyThrGlnGluTyr 600
Db 1741 AGCCAGTACAATGAAGACATGAAGAAGTTTAGGAAAATGGCACTTTGAACCTCAAGAGTAC 1800
Qy 601 AsnAlaThrGlyGluTyrSerAsnProThrSerAspTyrGlyLeuTyrProSerGlySer 620
Db 1801 AACGCCAGGGTGAGTACAGTAATCCGACCAGTGACTATGGACTGTATCCCGTCTGGTTCA 1860
Qy 621 SerSerGluGlyGlnThrThrArgGluMetGluMetGlyLysIleLysArgThrGlyGln 640
Db 1861 AGCAGGAGGCCCAACACACCGCAATGGAGATGGGGAAGATTAAAGAACCCGGTCAG 1920
Qy 641 GlyTyrSerGlyProSerLeu 647
Db 1921 GGTATAGTGGACCTTCTCTT 1941

RESULT 3

AX825705 LOCUS 2189 bp DNA linear PAT 11-DEC-2003
DEFINITION Sequence 3 from Patent WO03072763.
ACCESSION AX825705

VERSION AX825705.1 GI:39751232

KEYWORDS Brassica napus (rape)

SOURCE Brassica napus

ORGANISM Brassica napus

REFERENCE 1 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

AUTHORS Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

TITLE rosid; eurosids II; Brassicales; Brassicaceae; Brassica.

JOURNAL Goring, D., Silva, N. and Haffani, Y. Z.

Patent: WO 03072763-A 3 04-SEP-2003;

Goring, Daphne (CA) ; Silva, Nancy (CA) ; Haffani, Yosr, Z. (CA)

FEATURES Location/Qualifiers

source 1..2189

/organism="Brassica napus"

/mol_type="unassigned DNA"

/db_xref="taxon:3708"

ORIGIN

Alignment Scores:

Pred. No.: 4,39e-76 Length: 2189

Score: 3453.00 Matches: 647

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 6 Gaps: 0

US-10-086-464-2 (1-647) x AX825705 (1-2189)

Qy 1 MetSerAlaProSerProGlyThrGlySerProProSerProProSerAsnSerThr 20

Db 97 ATGTCCTCGCGCGCTCTCGGGGAGCTGGTTCGCTCCATCTCCACCATCAATCCACA 156

Qy 21 ThrThrThrProProAlaSerAlaProProProThrProSerSerProProPro 40

Db 157 ACCACCACTCTCTCCAGCTTCGCTCTCTCCACCAACCTCTCTCTCTCTCTCTCT 216

Qy 41 ProSerThrIleProThrSerProProSerSerArgSerThrProSerAlaProPro 60

Db 217 CCATCCACTATTCGACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 276

Qy 61 ProSerProThrProSerThrProGlySerProProProProProGlnProSerPro 80

Db 277 CCATCTCCACCACTCCATCTACCGCGGATCTCCACCTCTCTCTCTCTCTCTCTCA 336

Qy 81 ProAlaProThrThrProGlySerProProAlaProValThrProProThrArgAsnPro 100

Db 337 CCCGCTCAACTACGCCCGGATCTCCACCGCGACCTGTACTCTCTCTCTCTCTCTCT 396

Qy 101 ProProSerValProGlyProProSerAsnProSerArgGluGlyGlySerProArgPro 120

Db 397 CCACCTTCAGTCCAGGACCACTCCCAATCTTCCGGAAGGAGATCTCTCTCTCTCT 456

Qy 121 ProSerSerProProSerProSerSerAspGlyLeuSerThrGlyValValVal 140

Db 457 CCATCT 516

Qy 141 GlyIleAlaIleGlyValAlaLeuLeuValIleValThrLeuIleCysLeuLeuCys 160

Db 517 GGAATCGCCATCGGAGGAGTCGCTCTGCTTGTGATAGTACTCTGATTGTCTCTCTCT 576

Qy 161 LysLysLysArgArgArgAspGluAspAlaTyrTyrValProProProProProPro 180

Db 577 AAGAAGAAACACGCGAGAGAGAGATGCTTACTATGTCTCTCTCTCTCTCTCTCTCTCT 636

Qy 181 GlyProLysAlaGlyGlyProTyrGlyGlyGlnGlnGlnGlnGlnGlnGlnGlnGln 200

DB:	8	Gaps:	0
US-10-086-464-2 (1-647) x AY028699 (1-2189)			
Qy	1	MetSerSerAlaProSerProGlyThrGlySerProProSerProProSerAsnSerThr	20
Db	97	ATGTCCTCGGCGCGCTCTCGGGGACTGGTTCCGCTCCATCTCCACCATCAACTCCACA	156
Qy	21	ThrThrThrProProProAlaSerAlaProProProThrThrProProSerSerProPro	40
Db	157	ACCACCACTCTCTCTCCAGCTCCGCTCTCTCTCCACACACACCTCTCTCTCTCCGCGC	216
Qy	41	ProSerThrIleProThrSerProProSerProSerSerArgSerThrProSerAlaProPro	60
Db	217	CCATCCACTATTCGACATCT	276
Qy	61	ProSerProProThrProSerThrProGlySerProProProProLeuProGlnProSerPro	80
Db	277	CCATCTCCACCACTCCATCTACGCGGGGATCTCCACCTCTCTCTCTCTCTCTCTCTCTCA	336
Qy	81	ProAlaProThrThrProGlySerProProAlaProValThrProProThrArgAsnPro	100
Db	337	CCCGCTCCAACTACGCGCGGATCTCCACCGCACCTGTACTCTCTCTCTCTCTCTCTCTCT	396
Qy	101	ProProSerValProGlyProProProSerAsnProSerArgGluGlyGlySerProArgPro	120
Db	397	CCACCTTCAGTCCCAAGACCAACCGTCCAAATCTTCCACGGAAGGAGGATCTCTCGACCT	456
Qy	121	ProSerSerProSerProProSerProSerSerAspGlyLeuSerThrGlyValValVal	140
Db	457	CCATCTCTCTCTCGCGCGCT	516
Qy	141	GlyIleAlaIleGlyGlyValAlaAlaLeuLeuValIleValThrLeuIleCysLeuLeuCys	160
Db	517	GGAAATCGCCATCGGAGGAGTCTGCTCTGTGTAGTAGTACTCTGATTGTCTCTCTCTGT	576
Qy	161	LysLysLysArgArgArgAspGluAspAlaTyrTyrValProProProProProProPro	180
Db	577	AAGAAGAAACACGCGAGAGAGAGAAATGCTTACTACTATGTCTCTCTCTCTCTCTCTCT	636
Qy	181	GlyProLysAlaGlyGlyProTyrGlyGlyGlnGlnGlnGlnTyrArgGlnGlnAlaAla	200
Db	637	GGTCCCAAGCCGGAGGACCTTACGTTGGACAGCAGCAACAATGGCGGCAACAAACGCA	696
Qy	201	ThrProProSerAspHisValValThrSerLeuProProProProProProProProPro	220
Db	697	ACACCACCGTCAGATCATGTGTGACGTCACCTTACCACCAACCACTTAAGGCTCCATCTCA	756
Qy	221	ProArgGlnProProProProProProProProPheMetSerSerSerGlyGlySerAsp	240
Db	757	CCACGGCAACCTCTCTCCACCTCCACCGCTTTTCATGAGCAGCGCGGCTCCGAC	816
Qy	241	TyrSerAspArgProValLeuProProProSerProGlyLeuValLeuGlyPheSerLys	260
Db	817	TACTCGGACCGTCCAGTCTTCTCTCCCGCTCTCCAGGCTTGTGTAGGCTTCTCCAAA	876
Qy	261	SerThrPheThrTyrGluGluLeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeu	280
Db	877	AGCACTTTCCATACAGGAGCTAGTAGCCACCAATGGTTTCTCCGAGGCGAACTTG	936
Qy	281	LeuGlyGlnGlyGlyPheGlyTyrValHisLysGlyValLeuProSerGlyLysGluVal	300
Db	937	TTAGGACAAGCGGGTTCGGTTACGTGCACAAAGGTGTGTTCCTTAGTGGGAAAGAAAGTT	996
Qy	301	AlaValLysGlnLeuLysValGlySerGlyGlnGlyGluArgGluPheGlnAlaGluVal	320
Db	997	GCTGTGAACAGTGTGAAGATTTGGGAGTGTTCAGGGAGAGAGGAGTTTCAGCAGAGTT	1056
Qy	321	GluIleIleSerArgValHisHisArgHisLeuValSerLeuValGlyTyrCysIleAla	340
Db	1057	GAGATCATCAGCAGAGTTCACACAGGCATCTGGTGTCTCTGTGTGTATTGATTCGCC	1116
Qy	341	GlyAlaLysArgLeuLeuValTyrGluPheValProAsnAsnLeuLeuLeuHisLeu	360
Db	1117	GGTCCAAAGATTTGCTTCTATGAGTTTGTCTTAACAACAATCTCGAGCTTCACCTC	1176
Qy	361	HisGlyGluGlyArgProThrMetGluTrpSerThrArgLeuLysIleAlaLeuGlySer	380
Db	1177	CATGGCAGGAGCGCTCAATGGAATGGAGCACCAAGATTGAAGATTGCTCTTTGGATCT	1236
Qy	381	AlalysGlyLeuSerTyrLeuHisGluAspCysAsnProLysIleIleHisArgAspIle	400
Db	1237	GCTAAAGGACTTCTTATCTTCATGAAGATTGCAATCCTAAATCATTTACCGGTATATC	1296
Qy	401	LysAlaSerAsnIleLeuIleAspPheLysPheGluAlaLysValAlaAspPheGlyLeu	420
Db	1297	AAGGCTTCAACATATTGATAGATTTCAGTTTGAAGCTAAGGTTCGTATTTGGTCTT	1356
Qy	421	AlalysIleAlaSerAspThrAsnThrHisValSerThrArgValMetGlyThrPheGly	440
Db	1357	GCTAAGATTGCTTCTGATACAAACACGATGATCAACACGCTGTGATGGGAACCTTTGGG	1416
Qy	441	TyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPheSer	460
Db	1417	TACTTGGCTCGGAATACGCTGCAAGCGGAAGCTCACGGAGAAGTCTGACGTTTCTCA	1476
Qy	461	PheGlyValValLeuLeuLeuIleThrGlyArgArgProValAspAlaAsnVal	480
Db	1477	TTTGGCGTTGCTTTTGGAGCTCATTTACTGGAGCTGACCCGTTGATGCCACAATGTC	1536
Qy	481	TyrValAspAspSerLeuValAspTrpAlaArgProLeuLeuAsnArgAlaSerGluGln	500
Db	1537	TATGTAGATGACAGCTTAGTTGACTGGGCACGACCATTTGCTTAAACCGAGCATCTGAGCAA	1596
Qy	501	GlyAspPheGluGlyLeuAlaAspAlaLysMetAsnAsnGlyTyrAspArgGluGluMet	520
Db	1597	GGAGACTTTGAGGGTTTAGCTGATGCAAGAGTAAATAATGGGTATGACAGAGAGAGATG	1656
Qy	521	AlaArgMetValAlaCysAlaAlaAlaCysValArgHisSerAlaArgArgProArg	540
Db	1657	GCTCCGATGGTTGCTTGTGCTGGCGCTGTGTTCGCCATTCAGCTCGCCGAGACCTCGC	1716
Qy	541	MetSerGlnIleValArgAlaLeuGluGlyAsnValSerLeuSerAspLeuAsnGluGly	560
Db	1717	ATGAGCCAGATTGTCGCTGCGTTAGAAGGAAATGTATCACTGTCTGATCTTTAAGCAAGGG	1776
Qy	561	MetArgProGlyGlnSerAsnValTyrSerSerTyrGlyGlySerThrAspTyrAspSer	580
Db	1777	ATGAGACAGGTCAAGCAATGTATACAGCTCATACGGAGGAAGACCGATTATGACTCG	1836
Qy	581	SerGlnTyrAsnGluAspMetLysLysPheArgLysMetAlaLeuGlyThrGlnGluTyr	600
Db	1837	AGCCAGTACATGAAGACATGAAGAAGTTTAGGAAAATGGCACTTGGAACTCAAGAGTAC	1896
Qy	601	AsnAlaThrGlyGlyTyrSerAsnProThrSerAspTyrGlyLeuTyrProSerGlySer	620
Db	1897	AACGCCACGGGTAGTACAGTAATCCGACCACTATGAGCTGATGACCTGATGACCTGCTTCA	1956
Qy	621	SerSerGluGlyGlnThrThrArgGluMetGluMetGlyLysIleLysArgThrGlyGln	640
Db	1957	AGCAGCAGGCGCCAAACCAACACCGGAAATGGAGATGGGGAAGATTAAAGAAACCGGTGAG	2016
Qy	641	GlyTyrSerGlyProSerLeu	647
Db	2017	GGTTATAGTGACCTTCTCTT	2037
RESULT 5			
AX825738			
LOCUS			
DEFINITION			
SEQUENCE 36 from Patent WO03072763.			
AX825738			
ACCESSION			
VERSION			
AX825738.1			
GI:39751255			
KEYWORDS			
Arabidopsis thaliana (thale cress)			
Arabidopsis thaliana			
SOURCE			
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE

1 Goring, D., Silva, N. and Haffani, Y. Z.
Increasing plant seed production
Patent: WO 03072763-A 36 04-Sep-2003;
JOURNAL Daphne (CA) ; Silva, Nancy (CA) ; Haffani, Yosr, Z. (CA)
Goring, Location/Qualifiers
FEATURES
1. 1959
source /organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"

ORIGIN

Alignment Scores:
Pred. No.: 1,23e-63 Length: 1959
Score: 2936.50 Matches: 574
Percent Similarity: 90.61% Conservative: 24
Best Local Similarity: 86.97% Mismatches: 41
Query Match: 85.04% Indels: 21
DB: 6 Gaps: 12

US-10-086-464-2 (1-647) x AX825738 (1-1959)

```
Qy 1 MetSerSerAlaProSerProGlyThr-----GlySerProProSerProProSer 17
Db 1 ATGTCACAGCGCGTCTCCAGGCACCTACTCCATCACCCTCCACCGTCTCTCCACCA 60
Qy 18 AsnSerThrThrThrProProProProLa-----SerAlaProProThrThrPro 35
Db 61 AACTCGCAACACCACTCTCTCCAGCACCTCTCTCTCTCTCTCTCTCTCTCTCTCT 120
Qy 36 SerSerProProPro-----ProSerThrThrThrThrThrThrThrThrThrThr 54
Db 121 TCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180
Qy 55 ThrProSerAlaProProProSerProProThrProSerThrProGlySerProPro 74
Db 181 CCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 228
Qy 75 LeuProGlnProSerProProAlaProThrThrProGlySerProProAlaProValThr 94
Db 229 CTTCCTCAACCTCTCCCGCTCCGCTCCATCACTCTCT---TCTCCACCGTCTCCACCA 285
Qy 95 ProProThr---ArgAsnProProSerValProGlyProProProSerAsnProSerArg 113
Db 286 CCCTCAAAACCTCGAAGCCCTCCATCTCTCTTAACCAAGGACCAACCACTCTCCCTCA 342
Qy 114 GluGlyGlySerProArgProProSerSer-----ProSerProProSerProSerSer 131
Db 343 ---GGATCTACTCTAGAACCTCCATCAACACACTAAACCGTCGCGCGTCTGTATTCTCC 399
Qy 132 AspGlyLeuSerThrGlyValValValGlyLeuAlaLeuGlyValAlaLeuVal 151
Db 400 GATGGATTGTCTACCGGAGTTGTGTAGGATCGCCATTTGGAGGAGTCTGTTCTTGT 459
Qy 152 IleValThrLeuIleCysLeuLeuCysLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 169
Db 460 ATACTGACTCTGATTGTCTCTCTCTGTAAGAAGAAAGAAAGAAAGAAAGAAAGAA 519
Qy 170 AspAlaTyrTyrValProProProProProProGlyProGlyProGlyProGly 189
Db 520 GCTGTTTACTATGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 579
Qy 190 GlyGlnGlnGlnThrArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 208
Db 580 GGTCAACACAGTATTGGCAACAAACACACACACACACACACACACACACACACAC 639
Qy 209 ThrSerLeuProProProProProProProProProProProProProProProPro 228
Db 640 ACGTCATTG---CCACACACCTTAAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 696
Qy 229 ProProPro-----PheMetSerSerSerGlyGlySerAspTyrSerAspArgProValLeu 247
```

```
Db 697 CCACCACCAAGCAATTCATGAGTAGCAGTGGTGGTCTCTGACTATTCCGATCTTCCGGTCTT 756
Qy 248 ProProProSerProGlyLeuValLeuGlyPheSerLySerThrPheThrTyrGluGlu 267
Db 757 CTCCACCACTCTCCAGGCTTGTGTAGGCTTTTCTAAAGACACTTTTCACTTATGAGAG 816
Qy 268 LeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyGlyPheGly 287
Db 817 TTGTCGAGAGCTACTAATGCTCTCTGAGGCTAATTTGTTAGGACAAGGAGGCTTGGT 876
Qy 288 TyrValHisLyGlyValLeuProSerGlyLyGluValAlaValLyGlnLeuValVal 307
Db 877 TATGTGCATAAAGGTATATTCCTAGTGGGAAAGAGTCTGTGAAACAGTTCGAAAGCT 936
Qy 308 GlySerGlyGlnGlyGluArgGluPheGlnAlaGluValGluLeuIleLeuSerArgValHis 327
Db 937 GGTAGTGGTTCAGGAGAGAGAGAGTTCAGGCTGAGGTTGAGATCATTAGCAGAGTTCAT 996
Qy 328 HisArgHisLeuValSerLeuValGlyTyrCysIleAlaGlyAlaLyArgLeuValVal 347
Db 997 CACAGGCAATTTGGTTCCTTATTTGGTTATTTGATGCGCGTGTTCGAAAGATTACTTGT 1056
Qy 348 TyrGluPheValProAsnAsnLeuGluLeuHisLeuHisGlyGlyGlyArgProThr 367
Db 1057 TATGAGTTTCTTCCAAACAACAAATCTTGAGTTTCCCTCCATGTTAAGGACGCGCTACG 1116
Qy 368 MetGluTrpSerThrArgLeuLeuValLeuAlaGlySerAlaLyGlyLeuSerTyrLeu 387
Db 1117 ATGGAATGGAGTACTAGATTGAAGATTGCTCTTGATCTCTCTAAGGACATTCATATCTT 1176
Qy 388 HisGluAspCysAsnProLyIleLeuHisArgAspIleLyAlaSerAsnIleLeuIle 407
Db 1177 CATGAAGATTGCAATCCGAAATCATTCACCGTGATATTAAAGCGTCAACATATTGATT 1236
Qy 408 AspPheLyGlyGluAlaLyValAlaAspPheGlyLeuAlaLyIleAlaSerAspThr 427
Db 1237 GATTTCAAATTTGAAGCTAAGGTTGCTGACTTTGGTCTTGCCAAAGATTGCTCTGTATACA 1296
Qy 428 AsnThrHisValSerThrArgValMetGlyThrPheGlyTyrLeuAlaProGlyTyrAla 447
Db 1297 AACACTCATGTATCTACACGCTGATGGGAACTTTGGGTATTGGCTCCGGAATATGCT 1356
Qy 448 AlaSerGlyLyGlyLeuThrGluLySerSerAspValPheSerPheGlyValValLeuLeuGlu 467
Db 1357 GCAGTGGAAAGCTCACAGAAAGTCTGACGTTTCTCATTTGGCGTGTACTTTTGAA 1416
Qy 468 LeuIleThrGlyArgArgProValAspAlaAsnAsnValTyrValAspAspSerLeuVal 487
Db 1417 CTTATTACTGGAGGCGCCCTGTTGATGCGAACAAATGCTATGTATGATGACAGCTTAGTT 1476
Qy 488 AspTrpAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGlyLeuAla 507
Db 1477 GACTGGGACAGCATTTGCTTAAACCGGACATCTGAGAGAGAGATGGCTCGCATGGTGTGCT 1536
Qy 508 AspAlaLyGlyMetAsnAsnGlyTyrAspArgGluGluMetAlaArgMetValAlaCysAla 527
Db 1537 GATTCAAAGATGGGTAAATGAGTATGACAGAGAGAGATGGCTCGCATGGTGTGCTGGCT 1596
Qy 528 AlaAlaCysValArgHisSerAlaArgArgArgProArgMetSerGlnIleValArgAla 547
Db 1597 GCGGCTTGTCTGCCATTGAGTCTCGCGCAGACCTCGCATGAGCCAGATAGTAGCGGCG 1656
Qy 548 LeuGluGlyAsnValSerLeuSerAspLeuAsnGluGlyMetArgProGlyGlnSerAsn 567
Db 1657 TTAGAAGGAAATGTATCGCTGTCTGATCTTAAAGGAGGATGAGACCGGGTCAAGCAAC 1716
Qy 568 ValTyrSerSerTyrGlyGlySerThrAspTyrAspSerSerGlnTyrAsnGluAspMet 587
Db 1717 GTATACAGCTCATATGGAGGAGACACAGACTATGACAGGACCCATACACAGCAGCATG 1776
Qy 588 LysLysPheArgLyGlyMetAlaLeuGlyThrGlnGluTyrAsnAlaThrGlyGlyTyrSer 607
```

Db 1777 ATAAAGTTTAGGAAATGGCTCTTGGAACCTCAAGATAAGCGCACCAACCGCGAGTACAGT 1836

Qy 608 AnProThrSerAspTyrGlyLeuTyrProSerGlySerSerSerGlyGlnThrThr 627

Db 1837 AATCCACCAAGTACTAGCGAGTGTACCGCTCTGGTTCAAGCAGTGAAGGTCAAGCCACA 1896

Qy 628 ArgGluMetGluMetGlyValIleIysArgThrGlyGlnGlyTyrSerGlyProSerIleu 647

Db 1897 CGAGAAATGGAGATGGGAAGATTAAAGAACCCGGTCAAGGTTATAGTGGACCTCTCTT 1956

RESULT 6

BT008400

LOCUS

Arabidopsis thaliana A3924550 gene, complete cds.

BT008400

VERSION

BT008400.1 GI:30725473

KEYWORDS

FLI CDNA.

SOURCE

Arabidopsis thaliana (thale cress)

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

1 (bases 1 to 1959)

Cheuk,R., Chen,H., Kim,C.J., Shinn,P., Bowser,L., Carninci,P.,

Dale,J.M., Hayaehizaki,Y., Hsuan,V.W., Ishida,J., Jones,T.,

Kamiya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lin,J., Miranda,M.,

Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L.,

Sakurai,T., Satou,M., Seki,M., Southwick,A., Toriumi,M., Wong,C.,

Wu,H.C., Yamada,K., Yu,G., Yuan,S., Shinozaki,K., Davis,R.W.,

Theologis,A. and Ecker,J.R.

Arabidopsis ORF clones

Unpublished

2 (bases 1 to 1959)

Cheuk,R., Chen,H., Kim,C.J., Shinn,P., Bowser,L., Carninci,P.,

Dale,J.M., Hayaehizaki,Y., Hsuan,V.W., Ishida,J., Jones,T.,

Kamiya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lin,J., Miranda,M.,

Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L.,

Sakurai,T., Satou,M., Seki,M., Southwick,A., Toriumi,M., Wong,C.,

Wu,H.C., Yamada,K., Yu,G., Yuan,S., Shinozaki,K., Davis,R.W.,

Theologis,A. and Ecker,J.R.

Direct Submission

Submitted (15-MAY-2003) Salk Institute Genomic Analysis Laboratory

(SIGNAL), Plant Biology Laboratory, The Salk Institute for

Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,

USA

The discrepancy does not affect the protein sequence.

RIKEN Genomic Sciences Center (GSC) members carried out the

collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN

CDS

1. .1959

/note="unknown protein"

/codon_start=1

/product="At3g24550"

/protein_id="AAP37759.1"

/db_xref="GI:3072547"

/translation="MSTAPSGVTTPSPSPPTNTTTTPPPAASPPPTTTSSPP

PSSTNSTSPSPSPPLPSPSPSGSLTPPLPQSPSPAPITPSPSPPTSPSPSP

PSNQSPNTSPSGTSPSTPSTPSPSPSDSLSTGVVVGIAIGGVAILVILILCL

LCKRRRRHDDAAAYVPPPPSGGKAGPYGGQQYQQQNASRSDNNHVTSLPP

PKPSPRPKPPPPPPPPAFMSSSGSDYSDLPVLPSPSPGLVLFSGSTFTYELSR

TNGSLGICWAGVQRLLVYFVNNLEFHLHGKGRPTMWSRTRKLKLAGSKGLYL

HEDCNPKLIHRDIKASNILIDFPEAKVDFGLAKIASDNTNTHVSTVMGTFGYLA

YASGKSTGSDVFSFGVILVELITGRPPVANNVYVDDSLVDWARLLNRSBGDF

EGLADKMGNEYDREEMVACAAACVRRSRRRPRMSQIVRALEGNVLSLDNEMG

RPGHSNVSYSGSTDYDTSQYNDMDIKFRKMLGTQEGYGTGTSNPTSDYGLYPSG

SSSEGOATREMENGKIKKTKGGYSGPL"

misc_difference 1920

/note="compared to genome"

/replace="t"

ORIGIN

Alignment Scores:

Pred. No.: 1.23e-63 Length: 1959

Score: 2936.50 Matches: 574

Percent Similarity: 90.61% Conservative: 24

Best Local Similarity: 86.97% Mismatches: 41

Query Match: 85.04% Indels: 21

DB: 8 Gaps: 12

US-10-086-464-2 (1-647) x BT008400 (1-1959)

Qy 1 MetSerAlaProSerProGlyThr-----GlySerProSerProSer 17

Db 1 ATGTCCACAGCGCGTCTCCAGGCACTACTCCATCCATCCACCGTCTCTCCACA 60

Qy 18 AsnSerThrThrThrProProAla-----SerAlaProProThrPro 35

Db 61 AACTCGAACACCACTCTCTCCAGCAGCCTCTCTCTCCACCACTCTCT 120

Qy 36 SerSerProProPro-----ProSerThrThrThrSerProProSerSerArgSer 54

Db 121 TCTTCTCTCTCGTCCGCGCTCAACTAATCACTCTCTCTCTCTCTCTCTTTA 180

Qy 55 ThrProSerAlaProProProSerProThrProThrProGlySerProProPro 74

Db 181 CCT 228

Qy 75 LeuProGlnProSerProProAlaProThrThrProGlySerProProAlaProValThr 94

Db 229 CTTCTCAACCTTCCCTCTCCGCTCCCATCTCTCTCTCTCTCTCTCTCTCT 285

Qy 95 ProProThr---ArgAsnProProProSerValProGlyProProSerAsnProSerArg 113

Db 286 CCCTCAACCTCGAAGCCCTCCATCTCTCAACCAAGGACCACCAACACCTCTCA 342

Qy 114 GluGlySerProArgProProSerSer-----ProSerProProSerProSerSer 131

Db 343 ---GGATCTACTCTAGAACTCCATCAAACTAAACCGTCGCGCGCTCTGATCTTCC 399

Qy 132 AspGlyLeuSerThrGlyValValGlyIleAlaIleGlyGlyValAlaLeuVal 151

Db 400 GATGATGTCTACCGAGTGTGTAGGAATCCCATTTGGAGGATCGCTATTCTTGT 459

Qy 152 IleValThrLeuIleCysLeuLeuCysIleIleIleIleIleIleIleIleIle 169

Db 460 ATACTGACTGATTTGTCTTCTCTGTAAAGAAACGAAGACGACGACGATGAA 519

Qy 170 AspAlaTyrTyrValProProProProProProProProProProProProPro 189

Db 520 GCTGCTTACTATGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 579

Qy 190 GlyGlnGlnGlnTrpArgGlnGlnAsnAlaThrProProSerAsp---HisValVal 208
 Db 580 GGTCAACAACAGTATTGGCAACAACAAACCGGTACCGCGCTCAGATAATCATGTAGTG 639
 Qy 209 ThrSerLeuProProProProProProProProProProProProProProProProPro 228
 Db 640 ACGTCATTG---CACACCACTTAAGCCTCCATCTCCACCACCAAAACCTCTCCGCCACCT 696
 Qy 229 ProProPro---PheMetSerSerGlyGlySerAspTyrSerAspArgProValLeu 247
 Db 697 CCACCACCAAGCATCATGATGATGAGTGGTGTCTGACATATTGGATCTTCGGTTCCT 756
 Qy 248 ProProProSerProGlyLeuValLeuGlyPheSerLysSerThrPheThrTyrGluGlu 267
 Db 757 CCTCCACCATCTCCAGGGCTGTGTAGCTTTTCTAAAGACACTTTTCACATTATGAGGAG 816
 Qy 268 LeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyGlyPheGly 287
 Db 817 TTGTCGAGAGTACTTAATGCGCTTCTCTGAGGCTAATTTGTTAGGACCAAGGAGGTTTGGT 876
 Qy 288 TyrValHisLysGlyValLeuProSerGlyLysGluValAlaValLysGlnLeuLysVal 307
 Db 877 TATGTGCATAAAGTATATTCCTAGTGGGAAAGAGTGTCTGTGAAACAGTTCAAGCT 936
 Qy 308 GlySerGlyGlnGlyGluArgGluPheGlnAlaGluValGluIleLeuSerArgValHis 327
 Db 937 GGTAGTGGTCAGGAGAGAGAGAGTTTCAGGCTCAGGTTGAGATCATTTAGCAGAGTTTCAT 996
 Qy 328 HisArgHisLeuValSerLeuValGlyTyrCysIleAlaGlyAlaLysArgLeuLeuVal 347
 Db 997 CACAGGCATTGGTTCCTTATTGGTTATTGTATGCGCGGTGTCAAGAGTATTACTTGTTC 1056
 Qy 348 TyrGluPheValProAsnAsnLeuGluLeuHisLeuHisGlyGluGlyArgProThr 367
 Db 1057 TATGAGTTGTTCCAAACAAACAACTTGTGAGTTTCACTCCATGTTAAGGAGCGCTACG 1116
 Qy 368 MetGluTrpSerThrArgLeuLysIleAlaLeuGlySerAlaLysGlyLeuSerTyrLeu 387
 Db 1117 ATGGAATGGAGTACTAGATTCAAGATTGCTCTGATCTGTCTAAAGGACTTTTCATATCTT 1176
 Qy 388 HisGluAspCysAsnProLysIleLeuHisArgAspIleLysAlaSerAsnIleLeu 407
 Db 1177 CATGAAGATTGCAATTCGAAATCATTCACCGTGTATTAAGGCGTCAAAACATATTGATT 1236
 Qy 408 AspPheLysPheGluAlaLysValAlaAspPheGlyLeuAlaLysIleAlaSerAspThr 427
 Db 1237 GATTTCAAATTGAAGCTAAGTTGCTGACTTTGGTCTTGCACAGATGCTCTCTGATACA 1296
 Qy 428 AsnThrHisValSerThrArgValMetGlyThrPheGlyTyrLeuAlaProGluTyrAla 447
 Db 1297 AACACTCATGTATCTACACGGCTGATGGGAACCTTTGGGTATTTGGCTCCGGAATATGCT 1356
 Qy 448 AlaSerGlyLysLeuThrGluLysSerAspValPheSerPheGlyValValLeuLeuGlu 467
 Db 1357 GCAAGTGGAAAGCTCACAGAAAGTCTGACGTTTCTCATTTGGCGTTGTACTTTGGAA 1416
 Qy 468 LeuIleThrGlyArgArgProValAspAlaAsnValTyrValAspSerLeuVal 487
 Db 1417 CTTATTACTGGAGCGCCCTTGTATGCGAACATGTCTAATGATGATGACAGCTTAGTT 1476
 Qy 488 AspTrpAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGluGlyLeuAla 507
 Db 1477 GACTGGGCACGACCATCTTAACCGACATCTCAGGAAGGAGATTTTGAGGTTTGGCT 1536
 Qy 508 AspAlaLysMetAsnAsnGlyTyrAspArgGluGluMetAlaArgMetValAlaCysAla 527
 Db 1537 GATTCAAGATGGGTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1596
 Qy 528 AlaAlaCysValArgHisSerAlaArgArgProArgMetSerGlnIleValArgAla 547
 Db 1597 GCGGCTGTGTGCGCATTCAGCTCGCGCAGACCTCGCATGAGCCAGATAGTACGGGCG 1656
 Qy 548 LeuGluGlyAsnValSerLeuSerAspLeuAsnGluGlyMetArgProGlyGlnSerAsn 567

Db 1657 TTAGAAGCAATGATCGCTGTCTGATCTTAACCAAGGAGTATGACCGGGTCAACGCAAC 1716
 Qy 568 ValTyrSerSerTyrGlyGlySerThrAspTyrAspSerSerGlnTyrAsnGluAspMet 587
 Db 1717 GTATACAGCTCATATGAGGAGGACAGACTATGACAGGAGCCCAATCAACGACGACATG 1776
 Qy 588 LysLysPheArgLysMetAlaLeuGlyThrGlnGluTyrAsnAlaThrGlyGlyTyrSer 607
 Db 1777 ATAAAGTTTAGGAAATGGCTCTTGGAACTCAAGATACGACCAACCGCGGAGTACAGT 1836
 Qy 608 AsnProThrSerAspTyrGlyLeuTyrProSerGlySerSerSerGluGlyGlnThrThr 627
 Db 1837 AATCCAAACAGTACTACGAGCTGTACCGCTCTGTTCAAGCAGTGAAGTCAAGCCACA 1896
 Qy 628 ArgGluMetGluMetGlyLysIleLysArgThrGlyGlnGlyTyrSerGlyProSerLeu 647
 Db 1897 CGAGAAATGAGATGGGAAAGATCAAGAAACCGGTCAAGTTATATAGTGGACCTCTCTT 1956

RESULT 7
 AV128792
 LOCUS
 DEFINITION Arabidopsis thaliana protein kinase-like protein mRNA, complete cds.
 ACCESSION AV128792
 VERSION AV128792.1 GI:22136227
 KEYWORDS FL1 CDNA.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 REFERENCE 1 (bases 1 to 2098)
 AUTHORS Tripp,M., Southwick,A., Karlin-Neumann,G., Nguyen,M., Miranda,M., Palm,C.J., Bowser,L., Jones,T., Banh,J., Carninci,P., Chen,H., Cheuk,R., Chung,M.K., Hayashizaki,Y., Ishida,J., Kamiya,A., Kawai,J., Kim,C., Lin,J., Liu,S.X., Narusaka,M., Pham,P.K., Sakano,H., Sakurai,T., Satou,M., Seki,M., Shinn,P., Yamada,K., Hayashizaki,Y., Ecker,J., Theologis,A. and Davis,R.W.
 TITLE Direct Submission
 JOURNAL Submitted (01-JUL-2002) DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA
 COMMENT e-mail for correspondence: arab@sequence.stanford.edu

The RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-Length cDNA'): Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGEC (SSP) Consortium members constructed and sequenced the PUNI (ORF) clones using the RAFL cDNAs: Tripp,M., Nguyen,M., Southwick,A., Karlin-Neumann,G., Lam,B., Miranda,M., Palm,C.J., Bowser,L., Jones,T., Banh,J., Chen,H., Cheuk,R., Chung,M.K., Kim,C., Lin,J., Liu,S.X., Pham,P.K., Sakano,H., Shinn,P., Yamada,K., Ecker,J., Theologis,A. and Davis,R.W.

Tripp,M. (SSP/Stanford) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Davis,R.W. (SSP/Stanford) contributed equally to this work as PIs.

FEATURES
 source
 1..2098
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /clone="U16568"
 /ecotype="Columbia"
 /note="This clone is in pENTR/SD-dTopo. This is a cloned PCR product using RIKEN clone RAFL09-06-A17 (AY059901) as a template"
 1..1959
 /codon_start=1

CDS


```
Qy 588 LysLysPheArgLysMetAlaLeuGlyThrGlnGlnGlyTyrAsnAlaThrGlyGlyTyrSer 607
Db 1777 ATAAAGTTTAGGAAATCGCTTGGAACTCAAGATACGCACACCGCGAGTACAGT 1836

Qy 608 AenProThrSerAspTyrGlyLeuTyrProSerGlySerSerGlyGlnThrThr 627
Db 1837 AATCAACACGACTAGCGACTGACCGCTGCTGTTCAAGCAGTGAAGGTCAAGCCACA 1896

Qy 628 ArgGluMetGluMetGlyLysAlaLeuArgThrGlyGlnGlyTyrSerGlyProSerLeu 647
Db 1897 CGAGAAATGAGATCGGAAGATTAAAGAAACCGGTCAAGGTATTAGTGGACCTCTCTT 1956

RESULT 8
AY056788 2116 bp mRNA linear PLN 08-OCT-2001
LOCUS Arabidopsis thaliana AT3G24550/MOB24_8 complete cds.
DEFINITION Arabidopsis thaliana AT3G24550/MOB24_8
ACCESSION AY056788
VERSION AY056788.1 GI:15983764
KEYWORDS FLI CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1 (bases 1 to 2116)
Cheuk,R., Chen,H., Kim,C.J., Koesema,E., Meyers,M.C., Banh,J.,
Bower,L., Carninci,P., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y.,
Ishida,J., Jiang,P.X., Jones,T., Kamiya,A., Karlin-Neumann,G.,
Kawai,J., Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M.,
Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Pham,P.K.,
Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C.,
Toriumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K.,
Davis,R.W., Theologis,A. and Ecker,J.R.
Arabidopsis cdna clones
Unpublished
2 (bases 1 to 2116)
Cheuk,R., Chen,H., Kim,C.J., Koesema,E., Meyers,M.C., Banh,J.,
Bower,L., Carninci,P., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y.,
Ishida,J., Jiang,P.X., Jones,T., Kamiya,A., Karlin-Neumann,G.,
Kawai,J., Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M.,
Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Pham,P.K.,
Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C.,
Toriumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K.,
Davis,R.W., Theologis,A. and Ecker,J.R.
Direct Submission
Submitted (11-SEP-2001) Salk Institute Genomic Analysis Laboratory
(SIGNAL), Plant Biology Laboratory, The Salk Institute for
Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
USA
RIKEN Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAPL cDNAs (RAPL CDNA : 'RIKEN
Arabidopsis Full-Length cDNA') : Seki,M., Narusaka,M., Ishida,J.,
Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
Hayashizaki,Y. and Shinozaki,K.
The Salk, Stanford, PGSC (SSP) Consortium members carried out the
sequencing and annotation of the RAPL cDNAs: Cheuk,R., Chen,H.,
Kim,C.J., Koesema,E., Meyers,M.C., Shinn,P., Banh,J., Bower,L.,
Dale,J.M., Goldsmith,A.D., Jiang,P.X., Jones,T., Karlin-Neumann,G.,
Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M., Nguyen,M.,
Onodera,C.S., Palm,C.J., Pham,P.K., Quach,H.L., Southwick,A.,
Tang,C.C., Toriumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S.,
Davis,R.W., Theologis,A., and Ecker,J.R.
Cheuk,R. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to
this work. Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk)
contributed equally to this work as PIs.
Location/Qualifiers
1..2116
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"

FEATURES
source
US-10-086-464-2 (1-647) x AY056788 (1-2116)
ORIGIN
Alignment Scores:
Pred. No.: 1,3e-63 Length: 2116
Score: 2936.50 Matches: 574
Percent Similarity: 90.61% Conservative: 24
Best Local Similarity: 86.97% Mismatches: 41
Query Match: 85.04% Indels: 21
DB: 8 Gaps: 12
US-10-086-464-2 (1-647) x AY056788 (1-2116)
Qy 1 MetSerAlaProSerProGlyThr-----GlySerProProSerProProSer 17
Db 37 ATGTCCACAGCGCGTCTCCAGGCACCTCTCCATCCATCTCCACCGTCTCTCCACCA 96
Qy 18 AsnSerThrThrThrProProAla-----SerAlaProProThrThrPro 35
Db 97 AACTCGAACACCACTCTCTCTCCAGCAGCTCTTCTCTCTCCACCAACCTCT 156
Qy 36 SerSerProProPro-----ProSerThrIleProThrSerProProSerSerArgSer 54
Db 157 TCTTCTCTCTCTCGTCCGCTGCACTAATTAACCTCTCTCTCTCTCTCTCTCTT 216
Qy 55 ThrProSerAlaProProProSerProProThrProSerThrProGlySerProPro 74
Db 217 CCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 264
Qy 75 LeuProGlnProSerProProAlaProThrThrProGlySerProProAlaProValThr 94
Db 265 CTTCTCTCAACCTTCCCTCCGCTCCATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 321
Qy 95 ProProThr--ArgAsnProProSerValProGlyProProSerAsnProSerArg 113
Db 322 CCTCAACCTCGAAGCCCTCTCTCTTAACCAAGGACCAACCACTCTCTCTCTCTCT 378
Qy 114 GluGlyGlySerProArgProProProSerSer-----ProSerProProSerSer 131
Db 379 ---GGATCTACTCTAGACTCTCACTCAACACTAAACCGTCCGCGCTCTGATCTCTCC 435
Qy 132 AspGlyLeuSerThrGlyValValValGlyIleAlaIleGlyValAlaLeuLeuVal 151
Db 436 GATGGATTGTCTACCGGAGTTGGTAGGAATGCCATTGGAGGAGTCGCTATTCTTGT 495
Qy 152 IleValThrLeuIleCysLeuLeuCysLysLysLysArgArgArg-----AspGluGlu 169
Db 496 ATACTGACTCTGATTGCTCTCTCTGTAAGAGAAACGAGAGACGACGACGATGAA 555
Qy 170 AspAlaTyrTyrValProProProProProGlyProLysAlaGlyGlyProTyrGly 189
Db 170 AspAlaTyrTyrValProProProProProGlyProLysAlaGlyGlyProTyrGly 189
```


Db 556 GCTGCTTACTATGTTCTCTCTCCCTCCATCTGGTCCCAAGCTGGAGGACCTTACGGT 615
Qy 190 GlyGlnGlnGlnGlnTrpArgGlnGlnAenAlaThrProProSerAsp---HisValVal 208
Db 616 GGTCAACAACAGTATTGGCAACAACAAACCGTCCAGCGCGTCAGATAATCATGTAGTG 675
Qy 209 ThrSerLeuProProProProLysAlaProSerProProArgGlnProProProPro 228
Db 676 ACGTCATTG---CCACCACCTTAAGCCTCCATCTCCACCACCAAAACCTCCGCCACCT 732
Qy 229 ProProPro---PheMetSerSerGlyGlySerAspTyrSerAspArgProValLeu 247
Db 733 CCACCACAGCATTCATGATGACGATGGTGGTCTGACATATTCGGATCTTCGGTCTCT 792
Qy 248 ProProProSerProGlyLeuValLeuGlyPheSerLysSerThrPheThrTyrGluGlu 267
Db 793 CCTCCACCATCTCCAGGGCTGTGTAGGCTTTCTTAAAGACACTTTTCAATATGAGGAG 852
Qy 268 LeuAlaArgAlaThrAenGlyPheSerGluAlaAenLeuLeuGlyGlnGlyPheGly 287
Db 853 TTGTCGAGAGCTACTAATGGCTTCTCTGAGGCTAATTTGTTAGGACAAGGAGGGTTTGGT 912
Qy 288 TyrValHisLysGlyValLeuProSerGlyLysGluValAlaValLysGlnLeuLysVal 307
Db 913 TATGTGCATAAAGGTATATTGCTAGTGGGAAGAGTTGCTGTCAACACAGTTGAAAGCT 972
Qy 308 GlySerGlyGlnGlyGluArgGluPheGlnAlaGluValGluLeuLysLeuSerArgValHis 327
Db 973 GGTAGTGTGAGGAGAGAGAGAGTTTCAGCGCTGAGGTTGAGATCATTTAGCAGAGTTTCAT 1032
Qy 328 HisArgHisLeuValSerLeuValGlyTyrCysIleAlaGlyAlaLysArgLeuLeuVal 347
Db 1033 CACAGGCATTTGGTTTCTCTTATGGTTATTTGTTATGGCCGGTGTCAAGAGTATCTTGTG 1092
Qy 348 TyrGluPheValProAsnAsnLeuGluLeuHisLeuHisGlyGlyArgProThr 367
Db 1093 TATGATTTGTTTCCAAACCAACAATCTTGAGTTTCACTTCCATGTGTAAGGACGGCTACG 1152
Qy 368 MetGluTrpSerThrArgLeuLysIleAlaLeuGlySerAlaLysGlyLeuSerTyrLeu 387
Db 1153 ATGGAATGGAGTACTAGATTGAAGATTCTCTTGGATCTGTCTAAAGGACTTTCATATCTT 1212
Qy 388 HisGluAspCysAenProLysIleIleHisArgAspIleLysAlaSerAsnIleIle 407
Db 1213 CATGAAGATTGCAATCCCAAAATCAATCACCCTGATATTAAAGCGTCAAAATATTGATT 1272
Qy 408 AspPheLysPheGluAlaLysValAlaAspPheGlyLeuAlaLysIleAlaSerAspThr 427
Db 1273 GATTTCAATTTGAAGCTAAGGTTGCTGACTTTGGTCTTGGCAAGATTGCTTCTGATACA 1332
Qy 428 AenThrHisValSerThrArgValMetGlyThrPheGlyTyrLeuAlaProGluTyrAla 447
Db 1333 AACACTCATGTATCTACACGGCTGATGGGAACCTTTGGGTATTTGGCTCCGGAATATGCT 1392
Qy 448 AlaSerGlyLysLeuThrGluLysSerAspValPheSerPheGlyValValLeuLeuGlu 467
Db 1393 GCAAGTGAAGAGCTCACAGAAAGTCTGACGTTTCTCATTTGGCGTTGTACTTTTGGA 1452
Qy 468 LeuIleThrGlyArgArgProValAspAlaAenAenValTyrValAspSerLeuVal 487
Db 1453 CTTATTACTGGGAGCGCCCTGTTGATCGGACAAATGCTATGTATGATGACAGCTTAGTT 1512
Qy 488 AspTrpAlaArgProLeuLeuAenArgAlaSerGluGlnGlyAspPheGluGlyLeuAla 507
Db 1513 GACTGGGCACGACCATTCCTTAAACCGAGCATCTCAGGAAGGAGATTTTGGAGGGTTTGCT 1572
Qy 508 AspAlaLysMetAenAsnGlyTyrAspArgGluGluMetAlaArgMetValAlaCysAla 527
Db 1573 GATTCAAGATGGGTAATGAGTATGACAGAGAGGAGATGGCTCGCATGGTTGCTTGGCGCT 1632
Qy 528 AlaAlaCysValArgHisSerAlaArgArgArgProArgMetSerGlnIleValArgAla 547
Db 1633 GCGGCTTGTGTCGCATTCAGTTCGCGCAGACCTCGCATGAGCCAGATAGTACGGGCG 1692

Qy 548 LeuGluGlyAenValSerLeuSerAspLeuAsnGluGlyMetArgProGlyGlnSerAsn 567
Db 1693 TTAGAAGGAAATGATCGCTGTCTGATCTTAACGAAGGGATGAGACCGGGTCAGCAAC 1752
Qy 568 ValTyrSerSerTyrGlyGlySerThrAspTyrAspSerSerGlnTyrAsnGluAspMet 587
Db 1753 GTATACAGCTCATATGAGGAGACGACAGATATGACACGAGCCCAATACAGCAGCATG 1812
Qy 588 LysLysPheArgLysMetAlaLeuGlyThrGlnGluTyrAsnAlaThrGlyGluTyrSer 607
Db 1813 ATAAAGTTTAGGAAATGGCTCTTGGAACTCAAGATACGSCACAACCGCGAGTACAGT 1872
Qy 608 AsnProThrSerAspTyrGlyLeuTyrProSerGlySerSerGluGlyGlnThrThr 627
Db 1873 AATCCAAACGAGTACTACGAGCTGTACCCGCTCGTTCAGAGCATGAAGGTCACAGCCACA 1932
Qy 628 ArgGluMetGluMetGlyLysIleLysArgThrGlyGlnGlyTyrSerGlyProSerLeu 647
Db 1933 CGAGAAATGGAGATGGGAAGATTAAAGAAACCGGTCAAGGTTATATAGTGGACCTCTCTT 1992

RESULT 9

AY059901 2188 bp mRNA linear PLN 04-NOV-2001
LOCUS Arabidopsis thaliana protein kinase-like protein (MOB24.13) mRNA,
DEFINITION complete cds.

ACCSSION AY059901 GI:16649062

VERSION AY059901.1

KEYWORDS FLI CDNA.

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

1 (bases 1 to 2188)

NGUYEN, M., Karlin-Neumann, G., Southwick, A., Lam, B., Miranda, M.,
Palm, C. J., Bowser, L., Jones, T., Banh, J., Carninci, P., Chen, H.,
Cheuk, R., Chung, M. K., Hayashizaki, Y., Ishida, J., Kamiya, A.,
Kawai, J., Kim, C., Lin, J., Liu, S. X., Narusaka, M., Pham, P. K.,
Sakano, H., Sakurai, T., Satou, M., Seki, M., Shinn, P., Yamada, K.,
Shinozaki, K., Ecker, J., Theologis, A. and Davis, R. W.

Direct Submission

Submitted (22-OCT-2001) DNA Sequencing and Technology Center,
Stanford University, 855 California Avenue, Palo Alto, CA 94304,
USA

COMMENT e-mail for correspondence: arab@sequence.stanford.edu

RIKEN Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAFL cDNAs (RAFL CDNA : 'RIKEN
Arabidopsis Full-length cDNA'): Seki, M., Narusaka, M., Ishida, J.,
Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,
Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGEC (SSP) Consortium members carried out the
sequencing and annotation of the RAFL cDNAs: Nguyen, M.,
Southwick, A., Karlin-Neumann, G., Lam, B., Miranda, M., Palm, C. J.,
Bowser, L., Jones, T., Banh, J., Chen, H., Cheuk, R., Chung, M. K.,
Kim, C., Lin, J., Liu, S. X., Pham, P. K., Sakano, H., Shinn, P.,
Yamada, K., Ecker, J., Theologis, A. and Davis, R. W.

Nguyen, M. (SSP/Stanford) and Seki, M. (RIKEN GSC) contributed
equally to this work. Shinozaki, K. (RIKEN GSC) and Davis, R. W.
(SSP/Stanford) contributed equally to this work as PIs.

FEATURES
Location/Qualifiers

1..2188

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/db_xref="taxon:3702"

/chromosome="III"

/clone="RAFL09-06-A17"

/note="This clone is in pBluescript"

1..2188

/gene="MOB24.13"

gene

CDS

```
32. .1990
/gene="MOB24.13"
/codon_start=1
/product="protein kinase-like protein"
/db_xref="GI:16649063"
/db_xref="AAI24383.1"
/translation="MSTAPSGTTPSPSPPTNTSTTTTPPAASPPPTTTPSSPP
PSPSTNTPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP
PSNQGPNTPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP
LCKKRRRRHDEAAVYVPPPPSPGKAGGYGQQQYMQQNAARSPSDHNVTSLEP
PKPSPRPKPPPPPPAFMSSSGSDYSDLPVPPSPGLVGFSGSTFYIELSRA
HNGSEANLLGAGVGLVYVHKGILPSGKEVAVKQLKAGSGQGEREFQAEVEIISRVR
TLVSLIGYCMAGVGLVYVFPNNLEFHLHGKGRPTMESTRLKIALGSAKLSLV
HEDCNPKIHRDIKASNLIDFKPEAKVADFLAKIASDNTNTHVSTVMGTFFGLAPE
YAAGSKUTEKSDVFSFGVLLLELTGRPPVDANNVYDDSLVDWARPLLNARSEBGF
EGLADSKMGNEDYDEEMARMVACAAACVRHSARRPRMSQIVRALEGNVSLSDLNEM
RPHSHNVYSSVGGSTDDTDSQYNDMDMKFRKMLAQTEYGTGTEYSNPTSDYGLYPSG
SSSEGGQATREMENGKIKKTGGYSGPSL"
```

ORIGIN

Alignment Scores:

Pred. No.:	1,33e-63	Length:	2188
Score:	2936.50	Matches:	574
Percent Similarity:	90.61%	Conservative:	41
Best Local Similarity:	86.97%	Mismatches:	24
Query Match:	85.04%	Indels:	21
DB:	8	Gaps:	12

US-10-086-464-2 (1-647) x AV059901 (1-2188)

Qy	1	MetSerSerAlaProSerProGlyThr-----GlySerProProSerProSer 17	
Db	32	ATGTCACAGCGCGTCTCCAGGCACCTACTCCATCACCATCTCCACCGTCTCTCCCCACA 91	
Qy	18	AsnSerThrThrThrProProProAla-----SerAlaProProThrThrPro 35	
Db	92	AACTCGACAAACACCACTCTCTCCAGCAGCCCTCTCTCTCTCTCTCTCTCTCTCTCT 151	
Qy	36	SerSerProProPro-----ProSerThrThrThrThrThrThrThrThrThrThr 54	
Db	152	TCCT 211	
Qy	55	ThrProSerAlaProProProSerProThrProThrProThrProGlySerProPro 74	
Db	212	CCT 259	
Qy	75	LeuProGlnProSerProProAlaProThrThrThrProGlySerProProAlaProValThr 94	
Db	260	CTTCTCTCAACCTTCCCTCCGCTCCCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 316	
Qy	95	ProProThr---ArgAsnProProSerValProGlyProProSerAsnProSerArg 113	
Db	317	CCCTCAAAACCTCGAAGCCCTCCATCTCTCTCAACCAAGGACCAACCACTCTCCCTCA --- 373	
Qy	114	GluGlyGlySerProArgProProSerSer-----ProSerProProSerProSerSer 131	
Db	374	--GGATCTACTCTAGAACTCCATCAAACTAAACCGTCCGCGCTCTGTATCTCTCC 430	
Qy	132	AspGlyLeuSerThrGlyValValValGlyValGlyValGlyValGlyValGlyVal 151	
Db	431	GATGGATTGTCTACCGGAGTTGTGGTAGGAATCCCATCTGGTAGGAGTCTCTCTCTCT 490	
Qy	152	IleValThrLeuIleCysLeuLeuCysIleValValValValValValValValValVal 169	
Db	491	ATATCTGACTCTGATTGTCTCTCTCTGTAAGAAGAAACGAAGAAGACGACGATGAA 550	
Qy	170	AspAlaTyrThrValProProProProProGlyProGlyProGlyProGlyProGly 189	
Db	551	GCTGCTTACTATGTCT 610	
Qy	190	GlyGlnGlnGlnThrArgGlnGlnAsnAlaThrProProSerApp---HisValVal 208	
Db	611	GGTCAACAAAGTATTGGCAACAAACAAACCGCGTCACGCGCGTCAGATTAATCATGTAGTG 670	
Qy	209	ThrSerLeuProProProProLysAlaProSerProProArgGlnProProProProPro 228	
Db	671	ACGTCAATG---CCACCACTAGCCCTCCATCTCCACCAACCAACCTCTCTCCGCCACT 727	
Qy	229	ProProPro---PheMetSerSerSerGlyGlySerAspTyrSerAspArgProValLeu 247	
Db	728	CCACCACAGCATTCATGAGTAGCAGTGGTGTCTGCATATTTCGGATCTCTCCGGTCTT 787	
Qy	248	ProProProSerProGlyLeuValLeuGlyPheSerLysSerThrPheThrTyGluGlu 267	
Db	788	CCTCCACCATCTCCAGGCGTGTGTAGGCTTTCTTAAAGCATCTTCACATTATGAGGAG 847	
Qy	268	LeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyGlyPheGly 287	
Db	848	TTGTCGAGAGTACTAATGCTCTCTGAGGCTAATTTGTAGGCAAGGAGGTTTGGT 907	
Qy	288	TyrValHisLysGlyValLeuProSerGlyLysGluValAlaValLysGlnLeuLysVal 307	
Db	908	TATGTGCATAAAGTATATTGCTTAGTGGGAAAGTGTGCTGTGAAACAGTTGAAAGCT 967	
Qy	308	GlySerGlyGlnGlyGluArgGluPheGlnAlaGluValGluIleIleSerArgValHis 327	
Db	968	GGTAGTGGTCAGGAGAGAGAGAGTTCAGGCTCAGGTTGAGATCATTTAGCAGAGTTTCA 1027	
Qy	328	HisArgHisLeuValSerLeuValGlyTyrCysIleAlaGlyAlaLysArgLeuVal 347	
Db	1028	CACAGGCAATTTGGTCTCTTATTTGTTATGATGCGCGGTGTCAAGATTTACTTGTCT 1087	
Qy	348	TyrGluPheValProAsnAsnAsnLeuGluLeuHisLeuHisGlyGluGlyArgProThr 367	
Db	1088	TATGAGTTGTTTCCAAACAACATCTTGTAGTTTCCCTCCATGTTAAGGAGCGGCTACG 1147	
Qy	368	MetGluTrpSerThrArgLeuLysIleAlaLeuGlySerAlaLysGlyLysSerTyrLeu 387	
Db	1148	ATGGAATGGAGTACTAGATTGAAGATTTGCTTGGATCTGCTTAAAGGACTTTCTATATCT 1207	
Qy	388	HisGluAspCysAsnProLysIleIleHisArgAspIleLysAlaSerAsnIleLeu 407	
Db	1208	CATGAAGATTGCATCTCCGAAATCATTCACCGTGATTTAAGGCTCAACATATTGATT 1267	
Qy	408	AspPheLysPheGluAlaLysValAlaAspPheGlyLeuAlaLysIleAlaSerAspThr 427	
Db	1268	GATTTCAAATTTGAAGCTAAGTTGCTGACTTTGGTCTTGCACAAATTTGCTCTCTGATACA 1327	
Qy	428	AsnThrHisValSerThrArgValMetGlyThrPheGlyTyrLeuAlaProGluTyrAla 447	
Db	1328	AACACTCATGTATCTACACGCGTGATGGGAAACCTTTGGGTATTTGGCTCCGGAATATGCT 1387	
Qy	448	AlaSerGlyLysLeuThrGluLysSerAspValPheSerPheGlyValValLeuLeuGlu 467	
Db	1388	GCAAGTGGAAAGCTCACAGAAAGTCTGACGTTTCTCATTTGGCGTTGTACTTTTGGAA 1447	
Qy	468	LeuIleThrGlyArgArgProValAspAlaAsnAsnValTyrValAspAspSerLeuVal 487	
Db	1448	CTTATTACTGGAGGCGCCCTGTGTGATGCGAACAACTGCTATGTATGATGACAGCTTAGTT 1507	
Qy	488	AspTrpAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGlyLeuAla 507	
Db	1508	GACTGGGACGACCATTTGCTTAAACCGGATCTGAGGAAGGAGATTTTGGAGGTTTGGCT 1567	
Qy	508	AspAlaLysMetAsnAsnGlyTyrAspArgGluGluMetAlaArgMetValAlaCysAla 527	
Db	1568	GATTCAAAGATGGGTAAATGATGATGACAGAGAGAGATGCTCGCATGTGTGTTGCTTGGCT 1627	
Qy	528	AlaAlaCysValArgHisSerAlaArgArgProArgProMetSerGlnIleValArgAla 547	
Db	1628	CGGCTTGTGTTCGCCATTTCAGCTCGCGCAGACCTCGCATGAGCCAGATAGTAGCGGCG 1687	
Qy	548	LeuGluGlyAsnValSerLeuSerAspLeuAsnGluGlyMetArgProGlyGlnSerAsn 567	
Db	1688	TTAGAGGAAATGTATCGCTGTCTGATCTTTAAACGAAGGATGAGACCGGGTTCACAGCAAC 1747	

Qy 568 ValTyrSerSerTyrGlyGlySerThrAspTyrRhapSerSerGlnTyrAsnGluAspMet 587
 Db 1748 GTATACAGCTCATATGGAGGAGCACACAGCATATGACAGCGCCAAATCAACAGCAGCATG 1807
 Qy 588 LysIysPheAsgLysMetAlaLeuGlyThrGlnGluTyrAsnAlaThrGlyGluTyrSer 607
 Db 1808 ATAAAGTTTAGAAAATGGCTCTTGGAACTCAAGNATACGCACACACCGCGAGTACAGT 1867
 Qy 608 AsnProThrSerAspTyrGlyLeuTyrProSerGlySerSerSerGluGlyGlnThrThr 627
 Db 1868 AATCAACCACTGACTACGAGCTGTACCGCTCTGGTTCAAGCAGTGAAGGTCAAGCCACA 1927
 Qy 628 ArgGluMetGluMetGlyLysIleLysArgThrGlyGlnGlyTyrSerGlyProSerIleu 647
 Db 1928 CGAGAAATGGAGATGGGAAGATTAAAGAAACCGGTCAAGATTATAGTGGACCTCTCTT 1987

RESULT 10
 AY093065
 LOCUS AY093065 2190 bp mRNA linear PLN 21-APR-2002
 DEFINITION Arabidopsis thaliana unknown protein (At3g24550) mRNA, complete cds.
 ACCESSION AY093065
 VERSION AY093065.1 GI:20260331
 KEYWORDS FLI CDNA
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliopsida; eudicotyledons; Core eudicots; rosids; eucosids II; Brassicales; Brassicaceae; Arabidopsids.
 1 (bases 1 to 2190)
 Southwick,A., Karlin-Neumann,G., Nguyen,M., Lam,B., Miranda,M., Palm,C.J., Bowser,L., Jones,T., Banh,J., Carninci,P., Chen,H., Cheuk,R., Chung,M.K., Hayashizaki,Y., Ishida,J., Kamiya,A., Kawai,J., Kim,C., Lin,J., Liu,S.X., Narusaka,M., Pham,P.K., Sakano,H., Sakurai,T., Satou,M., Seki,M., Shinn,P., Yamada,K., Shinozaki,K., Ecker,J.J., Theologis,A. and Davis,R.W.
 Direct Submission
 Submitted (27-MAR-2002) DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA
 e-mail for correspondence: arab@sequence.stanford.edu

RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-length cDNA': Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGEC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Southwick,A., Nguyen,M., Karlin-Neumann,G., Lam,B., Miranda,M., Palm,C.J., Bowser,L., Jones,T., Banh,J., Chen,H., Cheuk,R., Chung,M.K., Kim,C., Lin,J., Liu,S.X., Pham,P.K., Sakano,H., Shinn,P., Yamada,K., Ecker,J.J., Theologis,A. and Davis,R.W.

Southwick,A., (SSP/Stanford) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Davis,R.W. (SSP/Stanford) contributed equally to this work as PIs.

FEATURES
 Location/Qualifiers
 1..2190
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /chromosome="3"
 /clone="RAFL07-12-E08 (R17522)"
 /ecotype="Columbia"
 /note="This clone is in pBluescript"

1..2190
 /gene="At3g24550"
 30..1988
 /gene="At3g24550"
 /codon_start=1
 /product="unknown protein"

/protein_id="AA013064.1"
 /cd_xref="GI:20260332"
 /translation="MSTAPSPGTTTPSPSPPTNSTTTTPPPAASPPPTTTSSPP
 PPSPTSTSPSPSPPLPSPSPGSLTPPSPSPAPITPSPSPPTPSPSPRSP
 PSNQPPNTPTSGSTPRTPSNTSPSPSDSGLSTGVVGIAGVAILVTLICL
 LCKKRRRHDDAAAYVPPPPSPGKAGPYGQOQYWOQONASRPDLNHHVTLSP
 PKPSPRKPPPPPPPPAFMSGGSDYSLPVLPPSPGLVLGFSKSTFTYBELSRA
 TNFSEANLLGQGGFGYVHGILPSSKEVAVKQLKAGSGQGERFQAEVILSRVHR
 HLVSILGYCMAGVQRLVLYEFVPPNNLEFHLHGKRPTEMWSTKLKIALSGLSYL
 HEDCNPKIIRHDKASNILDFKPAKVDVGLAKIASDTNTHVTMGTFFGYLAE
 YASNGKITEKSDVFSFGVILLELTIGRRPVADANNVYVDDSLVWARLNLRASEGF
 EGLASDMGNEYDREEMARVACAACVKKRRRRPMQSVLRALGNVSLSDLEMG
 RPDHSNVYSVGGSTDYDTQYNDMDMKRKMALGTQYGTGTSNPTSDYGLYPSG
 SSSEGOATREMENGKIKKTQGGSPSL"

ORIGIN

Alignment Scores:
 Pred. No.: 1.33e-63 Length: 2190
 Score: 2936.50 Matches: 574
 Percent Similarity: 90.61% Conservative: 24
 Best Local Similarity: 86.97% Mismatches: 41
 Query Match: 85.04% Indels: 21
 DB: 8 Gaps: 12
 US-10-086-464-2 (1-647) x AY093065 (1-2190)
 Qy 1 MetSerSerAlaProSerProGlyThr-----GlySerProSerProSer 17
 Db 30 ATGTCCACAGCGCGTCTCCAGGCACCTACTCATCATCTCCACCGTCTCTCCACCA 89
 Qy 18 AsnSerThrThrThrProProAla-----SerAlaProProThrPro 35
 Db 90 AACTCGAACACCACTCTCTCCAGCAGCCTCTTCTCTCTCTCCACCACTCTCT 149
 Qy 36 SerSerProProPro-----ProSerThrIleProThrSerProProSerSerArgSer 54
 Db 150 TCTTCTCTCTCTCGTCGCGGTCAACTTAATCAACCTCTCTCTCTCTCTCTCTCT 209
 Qy 55 ThrProSerAlaProProProSerProProThrProThrProGlySerProProPro 74
 Db 210 CCT 257
 Qy 75 LeuProGlnProSerProProAlaProThrThrProGlySerProProAlaProValThr 94
 Db 258 CTTCTCTCAACCTTCCCTCTCGCTCCCATCTACTCTCT---TCTCCACGCTCTCCACCA 314
 Qy 95 ProProThr---ArgAsnProProProSerValProGlyProProSerAsnProSerArg 113
 Db 315 CCCTCAACCTTCGAAGCCCTCCATCTCTTAACCAAGGACCACCAACACCTCCCTCA--- 371
 Qy 114 GluGlyGlySerProProProSerSer-----ProSerProProSerProSerSer 131
 Db 372 ---GGATCTACTCTAGAACTCCATCAAACTAAACCGTCGCGCGCTCTGATCTCTCC 428
 Qy 132 AspGlyLeuSerThrGlyValValGlyIleAlaIleGlyGlyValAlaLeuVal 151
 Db 429 GATGGATGTCTACCGAGTTGTGTAGAAATGCCATTCGAGGAGTGCCTATTCTTGT 488
 Qy 152 IleValThrLeuIleCysLeuLeuCysLysLysLysArgArgArg-----AspGluGlu 169
 Db 489 ATACTGACTCTGATTTGTCTTCTCTGTAAGAAAGAAACGAAGAAGACACGACGATGAA 548
 Qy 170 AspAlaTyrTyrValProProProProProGlyProLysAlaGlyGlyProTyrGly 189
 Db 549 GCTGCTTACTATGTTCCT 608
 Qy 190 GlyGlnGlnGlnGlnTyrArgGlnGlnAsnAlaThrProProSerAsp---HisValVal 208
 Db 609 GGTCAACAAACAGTATTGGCAACAAACAAACCGCTACCGCCCTCAGATAATCATGTAGT 668
 Qy 209 ThrSerLeuProProProProLysAlaProSerProProArgGlnProProProPro 228
 Db 669 ACGTCATTG---CCACCACTTAAGCCTTCTCTCCACCAAGAAACCTCTCTCTCTCTCT 725

Qy	229	ProProPro---PheMetSerSerSerGlyGlySerAspTyrSerAspArgProValLeu	247
Db	726	CCACCACCAGCATTCATGAGTAGCAGTGGTCTGACTATTCGGATCTCCGGTCTCT	785
Qy	248	ProProProSerProGlyLeuValLeuGlyPheSerIysSerThrPheThrTyrGluGlu	267
Db	786	CCTCCACCATCTCCAGGCGTGTGTAGGCTTTTCTAAAGACACTTTCACCTATGAGGAG	845
Qy	268	LeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyGlyPheGly	287
Db	846	TTGTTCGAGAGCTACTAATGGCTTCTCGAGGCTAATTTGTTAGCACAGGAGGGTTTGGT	905
Qy	288	TyrValHisIysGlyValLeuProSerSerGlyIysGluValAlaValIysGlnLeuIysVal	307
Db	906	TATGTGATAAAGGTATATTTCCTAGTGGGAAGAAGTGTGCTGTGAACACAGTTGAAAGCT	965
Qy	308	GlySerGlyGlnGlyGluArgGluPheGlnAlaGluValGluIleIleSerArgValHis	327
Db	966	GGTAGTGGTCAGGAGAGAGAGAGTTTCAGGCTGAGGTTGAGATCATTAGCAGAGTTTCAT	1025
Qy	328	HisArgHisLeuValSerLeuValGlyTyrCysIleAlaGlyAlaIysArgLeuLeuVal	347
Db	1026	CACAGGCATTTGGTTCTCTATTGGTTATTGTATGGCCGTGTTCAAAGATTACTTTGTC	1085
Qy	348	TyrGluPheValProAsnAsnAsnLeuLeuLeuHisIleuHisGlyGluGlyArgProThr	367
Db	1086	TATGAGTTTGTTCAAACCAACATCTTTGAGTTTCACCTCCATGTGTAAGGACGGCCTACG	1145
Qy	368	MetGluTrpSerThrArgLeuIysIleAlaLeuGlySerAlaIysGlyLeuSerTyrLeu	387
Db	1146	ATGGATGGAGTACTAGATTGAGATTGCTCTTGGATCTGCTAAAGACCTTTCATATCTT	1205
Qy	388	HisGluAspCysAsnProLyIleIleHisArgAspIleIysAlaSerAsnIleLeuIle	407
Db	1206	CATGAAGATTGCAATCCGAAATCATTCACCGTGATATTAAAGCGCTCAACATATTGATT	1265
Qy	408	AspPheIysPheGluAlaIysValAlaAspPheGlyLeuAlaIysIleAlaSerAspThr	427
Db	1266	GATTTCAAAATTTGAAGCTTAAGTTGCTGACTTTGGTCTTCCCAAGATTGCTTCTGATACA	1325
Qy	428	AsnThrHisValSerThrArgValMetGlyThrPheGlyTyrLeuAlaProGluTyrAla	447
Db	1326	AACACTCATGTATCTACACGGTGATGGGAACCTTTGGGTATTTGGCTCCGGAATATGCT	1385
Qy	448	AlaSerGlyIysLeuThrGluIysSerAspValPheSerPheGlyValValLeuLeuGlu	467
Db	1386	GCAAGTGGAAAGCTCACAGAAAGCTGACGTTTTCTCATTTGGCGTTGTACTTTTGGAA	1445
Qy	468	LeuIleThrGlyArgArgProValAspAlaAsnAsnValTyrValAspAspSerLeuVal	487
Db	1446	CTTATTACTGGAGCGCCCTGTTGATGCGCAACAATGCTATGTATGAGATGACAGCTTAGTT	1505
Qy	488	AspTrpAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGluGlyLeuAla	507
Db	1506	GACTGGGCACGACCATTTGCTTTAACCGAGCATCTGAGGAAGGAGATTTTGAAGGTTGGCT	1565
Qy	508	AspAlaIysMetAsnAsnGlyTyrAspArgGluGluMetAlaArgMetValAlaCysAla	527
Db	1566	GATTCAAAGATGGGTAATGAGTATGACAGAGAGAGATGGCTCGCATGGTTGCTTGGCT	1625
Qy	528	AlaAlaCysValArgHisSerAlaArgArgProArgMetSerGlnIleValArgAla	547
Db	1626	CGCGCTTGTGTTCGCCATTACGCTCGCCGACAGCTCGCATGAGCCAGATAGTACGGGCG	1685
Qy	548	LeuGluGlyAsnValSerLeuSerAspLeuAsnGluGlyMetArgProGlyGlnSerAsn	567
Db	1686	TTAGAAGGAATGTATCGCTGTCTGATCTTAAAGAAAGGATGACCGGGTTCACAGCAAC	1745
Qy	568	ValTyrSerSerTyrGlyGlySerThrAspTyrAspSerSerGlnTyrAsnGluAspMet	587
Db	1746	GTATACAGCTCATATGAGAAGACACAGATTATGACACGAGCCAAATCAACACGACGATG	1805

Qy 588 LysLysPheArgLysMetAlaLeuGlyThrGlnGlyTyraAsnAlaThrGlyGluTySer 607
 Db 1806 ATAAAGTTTATAGGAAAATGGCTCTTGTGAACTCTCAGAATAACGGCACAAACCGCGAGTACAGT 1865

Qy 608 AsnProThrSerAspTyTyGlyLeuTyProSerGlySerSerSerGluGlyGlnThrThr 627
 Db 1866 AATCCAAACCAAGTACGACGACTGTACCCGCTGCGTTCAAGCAGTGAAGTCAAGCCACA 1925

Qy 628 ArgGluMetGluMetGlyLysLysLysArgThrGlyGlnGlyTyTySerGlyProSerLeu 647
 Db 1926 CGAGAAATGAGATGGGAAGATTAAAGAAACCGGTCAAGTTATAGTGGACCTCTCTT 1985

RESULT 11
 BT008409 1959 bp mRNA linear PLN 15-MAY-2003
 LOCUS Arabidopsis thaliana At3g24600 gene, complete cds.
 DEFINITION
 ACCESSION BT008409
 VERSION BT008409.1 GI:30725491
 KEYWORDS FLI CDNA.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana

REFERENCE 1 (bases 1 to 1959)
 AUTHORS Cheuk,R., Chen,H., Kim,C.J., Shinn,P., Bowser,L., Carninci,P., Dale,J.M., Hayashizaki,Y., Hsuan,V.W., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Toriumi,M., Wong,C., Wu,H.C., Yamada,K., Yu,G., Yuan,S., Shinozaki,K., Davis,R.W., Theologis,A. and Ecker,J.R.
 TITLE Arabidopsis ORF clones
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1959)
 AUTHORS Cheuk,R., Chen,H., Kim,C.J., Shinn,P., Bowser,L., Carninci,P., Dale,J.M., Hayashizaki,Y., Hsuan,V.W., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Toriumi,M., Wong,C., Wu,H.C., Yamada,K., Yu,G., Yuan,S., Shinozaki,K., Davis,R.W., Theologis,A. and Ecker,J.R.
 TITLE Direct Submission
 JOURNAL Submitted (15-MAY-2003) Salk Institute Genomic Analysis Laboratory (SIGNAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA

COMMENT
 RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL CDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-length cDNA') : Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGEC (SSP) Consortium members constructed and sequenced the PUNI (ORF) clones using the RAFL CDNAs: Cheuk,R., Chen,H., Kim,C.J., Shinn,P., Bowser,L., Chan,M.M., Chang,C.M., Dale,J.M., Hsuan,V.W., Jones,T., Karlin-Neumann,G., Lam,B., Lee,J.M., Lin,J., Miranda,M., Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L., Southwick,A., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C., Yamada,K., Yu.G., Yuan,S., Davis,R.W., Theologis,A., and Ecker,J.R.

Cheuk,R. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk) contributed equally to this work as PIs.
 Location/Qualifiers
 1. .1959
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /chromosome="3"
 /clone="U15366"
 /ecotype="Columbia"

FEATURES
 source

Db 746 ACGTCATG---CCACCACCTAAGCCTCCATCTCCACCACGAAACCTCTCCGCCACCT 802

Qy 229 ProProPro---PheMetSerSerSerGlyGlySerAspTyrSerAspArgProValLeu 247

Db 803 CCACCACGACATTCATGAGTAGCAGTGGTGGTCTGACTATTCCGATCTTCCGGTCTT 862

Qy 248 ProProProSerProGlyLeuValLeuGlyPheSerLysSerThrPheThrThrGluGlu 267

Db 863 CCTCCACCATCTCCAGGGCTTGTGTAGGCTTTTCTAAAGACATTTTCACTTATGAGGAG 922

Qy 268 LeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyPheGly 287

Db 923 TTGTCGAGAGCTACTAATGGCTTCTCTGAGGCTAATTTGTAGCAAGAGGGTGTGGT 982

Qy 288 TyrValHisLysGlyValLeuProSerGlyLysGluValAlaValLysGlnLeuLysVal 307

Db 983 TATGTGCATAAAGGTATATTCCTAGTGGGAAAGAGTTGCTGTGAAACAGTTGAAAGCT 1042

Qy 308 GlySerGlyGlnGlyGluArgGluPheGlnAlaGluValGluLeuIleSerArgValHis 327

Db 1043 GGTAGTGGTCAGGAGAGAGAGATTTTCAGGCTGAGGTTGAGATCAATTAGCAGAGTTTCA 1102

Qy 328 HisArgHisLeuValSerLeuValGlyTyrCysIleAlaGlyAlaLysArgLeuVal 347

Db 1103 CACAGGCATTTGGTTTCTTTATTGGTTATTGTATGGCCGGTGTTCAAAGATTTACTTGT 1162

Qy 348 TyrGluPheValProAsnAsnLeuGluLeuHisLeuHisGlyGlyArgProThr 367

Db 1163 TATGAGTTTGTTCAAACCAACAATCTTGAGTTTCACTCCATGTTAAGGGAGCGCTACG 1222

Qy 368 MetGluTrpSerThrArgLeuValIleAlaLeuGlySerAlaLysGlyLeuSerTyrLeu 387

Db 1223 ATGGAATGGAGTACTAGATTGAAGATTCTCTTGGATCTGTCTAAAGGACTTTTCATATCT 1282

Qy 388 HisGluAspCysAsnProLysIleHisArgAspIleLysAlaSerAsnIleLeu 407

Db 1283 CATGAAGATTGCAATCCGAAATCATTCACCGTGATATTAAAGGGCTCAACATATTGATT 1342

Qy 408 AspPheLysPheGluAlaLysValAlaAspPheGlyLeuAlaLysIleAlaSerAspThr 427

Db 1343 GATTTCAAATTTGAAGCTAAGGTTGCTGACTTTGGTCTTGGCAAGATTGCTTCGTGATACA 1402

Qy 428 AsnThrHisValSerThrArgValMetGlyThrPheGlyTyrLeuAlaProGluTyrAla 447

Db 1403 AACACTCATGTATCTACACGGCTGATGGAAACCTTTGGGTATTGGCTCCGGAATATGCT 1462

Qy 448 AlaSerGlyLysLeuThrGluLysSerAspValPheSerPheGlyValValLeuLeuGlu 467

Db 1463 CCAAGTGGAAAGCTCACAGAAAGTCTGACGTTTCTCATTTGGCGTTGTACTTTTGGA 1522

Qy 468 LeuIleThrGlyArgArgProValAspAlaAsnValTyrValAspAspSerLeuVal 487

Db 1523 CTTATTACTGGAGGCGCCCTGTTGATCGCAACAATGCTATGTAGATGACAGCTTAGTT 1582

Qy 488 AspTrpAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGluGlyLeuAla 507

Db 1583 GACTGGGCACGACCATTCCTTAACCGACATCTGAGGAAGGAGATTTTGAAGGGTTGGCT 1642

Qy 508 AspAlaLysMetAsnAsnGlyTyrAspArgGluGluMetAlaArgMetValAlaCysAla 527

Db 1643 GATTCAAGATGGGTATGAGTATGACAGAGGAGGATGGCTCGCATGGTTGCTTGGCT 1702

Qy 528 AlaAlaCysValArgHisSerAlaArgArgProArgProArgMetSerGlnIleValArgAla 547

Db 1703 CGCGCTGTGTTCGCCATTACGTCGCCGACAGCCCTCGCATGAGCGCATAGTACGGGCG 1762

Qy 548 LeuGluGlyAsnValSerLeuSerAspLeuAsnGluGlyMetArgProGlyGlnSerAsn 567

Db 1763 TTAGAAGGAATGTATCGCTGTCTGATCTTAACGAAGGGATGAGACCGGGTCACAGCAAC 1822

Qy 568 ValTyrSerSerTyrGlySerThrAspTyrAspSerSerGlnTyrAsnGluAspMet 587

Db 1823 GTATACAGCTCATATGGAGGAAGCACAGACTATGACAGGCGCAATACACAGCAGCATG 1882

Qy 588 LysLysPheArgLysMetAlaLeuGlyThrGlnGlyTyrAsnAlaThrGlyGlySer 607

Db 1883 ATAAAGTTTATAGGAATGGCTTTGGAACTCAAGAATACGACACCGCGGAGTACAGT 1942

Qy 608 AsnProThrSerAspTyrGlyLeuTyrProSerGlySerSerGlyGlyGlnThrThr 627

Db 1943 AATCCAACAGTACTACGAGCTGTACCGTCTGGTTCAGCAGTGAAGTCAAGCCACA 2002

Qy 628 ArgGluMetGluMetGlyLysIleLysArgThrGlyGlnGlyTyrSerGlyProSerLeu 647

Db 2003 CGAGAAATGGAGATGGGAAGATTAAAGAAACCGTCAAGGTTATAGTGGACCTCTCT 2062

RESULT 13

AY089024 2324 bp mRNA linear PLN 14-APR-2003
LOCUS Arabidopsis thaliana clone 17909 mRNA, complete sequence.

DEFINITION AY089024

ACCESSION AY089024

VERSION AY089024.1 GI:21407798

KEYWORDS Full cDNA.

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

1 (bases 1 to 2324)

Haas,B.J., Volkovsky,N., Town,C.D., Troukhan,M., Alexandrov,N.,

Feldmann,K.A., Flavelli,R.B., White,O. and Salzberg,S.L.

Full-length messenger RNA sequences greatly improve genome

annotation

Genome Biol. 3 (6), RESEARCH0029 (2002)

22088475

PUBMED 12093376

2 (bases 1 to 2324)

Brover,V., Troukhan,M., Alexandrov,N., Lu,Y.-P., Flavelli,R. and

Feldmann,K.

Full-length cDNA from Arabidopsis thaliana

Unpublished

3 (bases 1 to 2324)

Brover,V., Troukhan,M., Alexandrov,N., Lu,Y.-P., Flavelli,R. and

Feldmann,K.

Direct Submission

Submitted (11-MAR-2002) Ceres, Inc, 3007 Malibu Canyon Road,

Malibu, CA 90265, USA

This clone sequence is one of 5,000 Ceres full-length cDNAs made

available to TIGR and Genbank. The following quality assessment of

this set was done by comparison with known proteins: two percent of

the clones are estimated to be 5'-truncated; less than one percent

are 3'-truncated; approximately two percent represent alternative

splice variants, including unspliced introns and spliced exons; one

percent may contain premature stop codons; five percent may have

frame shifts in a coding region. A sequence is considered to be

5'-truncated if it lacks the translation initiation start (ATG). A

sequence is considered to be 3'-truncated if it lacks the

C-terminal end of the encoded protein. Please note that these cDNA

sequences are derived from the Ws or Laer ecotypes and therefore

may contain polymorphisms when compared to sequences from Col-0.

Genet carried out the library production and sequencing of the

full-length clones. Ceres, Inc. carried out the clustering of the

5' sequences, selection of clones, and sequence assembly.

FEATURES

source

1. 2324

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/db_xref="taxon:3702"

/clone="17909"

ORIGIN

Alignment Scores:

Pred. No.: 2,45e-62 Length: 2324

Score: 2885.00 Matches: 570

Percent Similarity: 90.30% Conservative: 26

Best Local Similarity: 86.36% Mismatches: 42

Query Match:	83.55%	Indels:	23
DB:	8	Gaps:	13
US-10-086-464-2 (1-647) x AY089024 (1-2324)			
Qy	1	MetSerSerAlaProSerProGlyThr-----GlySerProProSerProProSer	17
Db	105	ATGTCACAGCGCGCTCTCCAGGCACTACTCCATCACCACCTCCACCGCTCTCCCTCCACCA	164
Qy	18	AsnSerThrThrThrProProAla-----SerAlaProProProThrPro	35
Db	165	AACTCGACACACACCTCTCTCCAGCAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	224
Qy	36	SerSerProProPro-----ProSerThrIleProThrSerProProProSerSerArgSer	54
Db	225	TCTTCTCTCTCTCGCGCGTCAACTAATCAACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	284
Qy	55	ThrProSerAlaProProProSerProProThrProSerThrProGlySerProProPro	74
Db	285	CCT	332
Qy	75	LeuProGlnProSerProAlaProThrProGlySerProProAlaProValThr	94
Db	333	CTTCTCTAACCTCTCTCACTCCGCTCCCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	389
Qy	95	ProProThr-----ArgAsnProProSerValProGlyProProSerAsnProSerArg	113
Db	390	CCCTCAAAACCTCGNAGCCCTCCATCTCTTAACCAAGGACCAACCAACCTCTCTCTCTCTCTCT	446
Qy	114	GluGlySerProArgProProSerSer-----ProSerProProProSerProSerSer	131
Db	447	---GGACTACTCTAGAACTCCATCAAAACGCTAAACCGTCCGCCACCGCTCTGATCTTCC	503
Qy	132	AspGlyLeuSerThrGlyValValValGlyIleAlaIleGlyGlyValAlaLeuVal	151
Db	504	GATGGATTGTCTACCGGAGTGTGGTAGGAATCGCCATTGGAGGAGTCTGATTTCTTGT	563
Qy	152	IleValThrLeuIleCysLeuLeuCysIleValValValValValValValValValVal	169
Db	564	ATACTGACTCTGATTTGTCTCTCTGTAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA	623
Qy	170	AspAlaTyrrValProProProProProProProGlyProGlyProGlyProGlyProGly	189
Db	624	GCTGCTTACTATGTT---CCT	680
Qy	190	GlyGln	208
Db	681	GGTCAACAAAGATATTGGCAACAAACAAACCGCTCACCGGCGTCAAGATAATCATGTAGTG	740
Qy	209	ThrSerLeuProProProProProProProProProProProProProProProProPro	228
Db	741	ACGTCAATG---CCACCACTTAAGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	797
Qy	229	ProProPro-----PheMetSerSerSerGlySerAspTyrrSerAspArgProValLeu	247
Db	798	CCACCACCCAGCATTCATGAGTAGCAGTGGTGGTCTGACTATTCCGATCTCTCTCTCTCTCTCT	857
Qy	248	ProProProSerProGlyLeuValLeuGlyPheSerIleSerIleSerIleSerIleSerIleSer	267
Db	858	CCTCCACCATCTCCAGGCTTGTGTAGGCTTTTCAAAAAGCACTTTTCACTTATGAGGAG	917
Qy	268	LeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyGlyPheGly	287
Db	918	TTGCAAGAGCTACTAATGGCTTCTCTGAGGCTAATTTGTAGGACAGGAGGTTTGGT	977
Qy	288	TyrValHisIleValLeuProSerGlyValValAlaValIleValIleValIleValIleVal	307
Db	978	TATGTGCATAAGGTATATTCCCTAGTGGGAAAGAGTGTGTGTGAAACAGTTGAAAGCT	1037
Qy	308	GlySerGlyGlnGlyGluArgGluPheGlnAlaGluValGluIleIleSerArgValHis	327
Db	1038	GGGAGTGTCTAGGAGAGAGAGATTTTCCAGCTGAGGTTGAGATCATTAGCAGAGTTTCAT	1097
Qy	328	HisArgHisLeuValSerLeuValGlyTyrrCysIleAlaGlyAlaGlyAlaGlyLeuVal	347
Db	1098	CACAGGCATTTGGTTTCTTTATTTGGTTATTTATGATGCTGATGTTCAAGATTTCTTGTCT	1157
Qy	348	TyrGluPheValProAsnAsnAsnLeuGluLeuHisLeuHisGlyGlyGlyArgProThr	367
Db	1158	TATGAGTTTGTCTCTCAACAACATCTTGGATTTTCCCTCCATGTTAAGGAGCCCTACAT	1217
Qy	368	MetGluTrpSerThrArgLeuValIleAlaLeuGlySerAlaIleGlyLeuSerTyrrLeu	387
Db	1218	ATGGAATGGAGTACTAGATTGAAGATTGCTCTTGGATCTGCTAAGGAGCTTTCATATCTT	1277
Qy	388	HisGluAspCysAsnProIleIleHisArgAspIleIleValIleValIleValIleValIle	407
Db	1278	CATGAAGATTGCAATCGAAATCATTCACCGTGATATTAAAGCGCTCAACATATTGATT	1337
Qy	408	AspPheIlePheGluAlaIleValAlaAspPheGlyLeuAlaIleIleIleIleIleIleIle	427
Db	1338	GATTTCAAATTTGAAGTAAAGTTGCTGACTTTGGTCTTCCAGGATTTGCTTCTGATACA	1397
Qy	428	AsnThrHisValSerThrArgValMetGlyThrPheGlyTyrrLeuAlaProGluTyrrAla	447
Db	1398	AACACTCATGTATCTACACCGGTGATGGGAACCTTTGGGTATTTGGCTCCAGATATGCT	1457
Qy	448	AlaSerGlyIleLeuThrGluIleValPheSerPheGlyValValLeuLeuGlu	467
Db	1458	GCAAGCGGAAAGCTCACAGAAAAGTCTGACGTTTCTCATTTGGTGTGTTGCTTTGGAG	1517
Qy	468	LeuIleThrGlyArgArgProValAspAlaAsnAsnValTyrrValAspAspSerLeuVal	487
Db	1518	CTTATAACCGGAGGCGCCCTGTTGATGCAACAATGCTATGTAGATGACAGCTTAGTT	1577
Qy	488	AspTrpAlaArgProLeuLeuAsnAlaAlaSerGluGlnGlyAspPheGlyGlyLeuAla	507
Db	1578	GACTGGGACAGCCCTTGTCTTAACCGGATCTGAGGAGGAGATTTTGATGTTGGCT	1637
Qy	508	AspAlaIleValMetAsnGlyTyrrAspArgGluGluMetAlaArgMetValAlaCysAla	527
Db	1638	GATTCAAAGATGGTAAATGAGTATAG-AGAGAGGAGATGCTCGCATGGTTGCTTGGCT	1696
Qy	528	AlaAlaCysValArgHisSerAlaArgArgArgProArgMetSerGlnIleValArgAla	547
Db	1697	GCAGCTTGTGTTGCCATTCAGCTCGCGCAGACCTCGCATGAGCCAGATAGTACGGCG	1756
Qy	548	LeuGluGlyValValSerLeuSerLeuAsnGluGlyMetArgProGlyGlnSerAsn	567
Db	1757	TTAAGAGAAATGTATCGCTGCTCTGATCTTAAAGAGGATGAGACCGGGTCAAGCAAC	1816
Qy	568	ValTyrrSerSerTyrrGlySerThrAspTyrrAspSerSerGlnTyrrAsnGluAspMet	587
Db	1817	GTATACAGCTCATATGGAGGAGCACTGACTATGACACGAGCAATACACGATGACATG	1876
Qy	588	LysIlePheArgLysMetAlaLeuGlyThrGlnGlyTyrrAsnAlaThrGlyGlyTyrrSer	607
Db	1877	AAGAAGTTTAGGAAATGGCTTTGGAATCTCAAGATACGCGCACACCGCGGATGACAGT	1936
Qy	608	AsnProThrSerAspTyrrGlyLeuTyrrProSerGlySerSerGlyGlnGlnThrThr	627
Db	1937	AATCCAAACAGTACACGAGCTGTACCGCTCTGGTCTCAAGCAGTGAAGTCAAGCCACA	1996
Qy	628	ArgGluMetGluMetGlyIleIleValArgThrGlyGlnGlyTyrrSerGlyProSerLeu	647
Db	1997	CGAGAAATGAGATGGGAAAGATTAGAAAACCGGTCAAGGTTATAGTGGACCTCTCTT	2056
RESULT 14			
AB020746			
LOCUS			
DEFINITION Arabidopsis thaliana genomic DNA, chromosome 3, P1 clone: MOB24.			
ACCESSION AB020746 BA000014			
VERSION AB020746.1 GI:3985949			
KEYWORDS Arabidopsis thaliana (thale cress)			
SOURCE Arabidopsis thaliana			
ORGANISM Arabidopsis thaliana			

REFERENCE AUTHORS TITLE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurooids II; Brassicales; Brassicaceae; Arabidopsis. 1 (sites) Kaneko,T., Katoh,T., Sato,S., Nakamura,A., Asamizu,E. and Tabata,S. Structural analysis of Arabidopsis thaliana chromosome 3. II. Sequence features of the 4,251,695 bp regions covered by 90 P1, TAC and BAC clones DNA Res. 7 (3), 217-221 (2000)
JOURNAL MEDLINE PUBMED	20363099
REFERENCE AUTHORS TITLE JOURNAL	10907853 2 (bases 1 to 79706) Kaneko,T., Kato,T., Sato,S., Nakamura,Y., Asamizu,E. and Tabata,S. Direct Submission Submitted (03-DEC-1998) Yaukazu Nakamura, Kazusa DNA Research Institute, Department of Plant Gene Research; 1532-3, Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:ynakamu@kazusa.or.jp, Tel:81-438-52-3935, Fax:81-438-52-3934) Address for correspondence: kaos@kazusa.or.jp For the latest information on annotation of this clone, please see http://www.kazusa.or.jp/kaos/cgi-bin/aggd.graph.cgi?c=MOB24 Genes with similarity to proteins in the databases are described in 'product' or 'note' qualifiers. Genes that have no significant protein similarity are described as 'unknown protein'. The software programs used to predict genes include: Grail (Informatics Group, Oak Ridge National Laboratory, http://complib.ornl.gov/Grail-1.3/), GENSCAN (Chris Burge, MIT, http://CCR-081.mit.edu/GENSCAN.html), NetGene2 (S.M. Hebsgaard, et al., CBS, Technical University of Denmark, http://www.cbs.dtu.dk/services/NetGene2/) and SplicePredictor (Volker Brendel, Stanford University, http://grainlin.zool.iastate.edu/cgi-bin/sp.cgi). Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, Washington University School of Medicine, St. Louis, http://genome.wustl.edu/eddy/tRNAscan-SE/). This sequence may not be the entire insert of this clone. It may be shorter because we remove overlaps between neighboring submissions. The 5' clone is MXP5 and the 3' clone is MSD24.
COMMENT	Location/Qualifiers 1..79706 /organism="Arabidopsis thaliana" /mol_type="genomic DNA" /strain="Columbia" /db_xref="taxon:3702" /chromosome="3" /clone_lib="Mitsui P1" 2398..-3399 /note="gene id:MOB24.1 unknown protein" /codon_start=1 /evidence=not_experimental /protein_id="BAB01995.1" /db_xref="GI:9294038" /translation="MGSEDETGYPKFYSLNRONHPMYSRPIPKRHAYNEEDED VEGEEDPQGYIIGNERFQRKPKNVSGFEAGSDAKVADNRQEAFVLELV GDRILQRRSLRNEWNAEKVSEELRSEETQCRMTDDIKRKREKI KVEKS GLSSKWSFFNKDMLLCVSPKSLGACGVDSGEFVPMNTKVLVDKSNQDMDSP GDSEEDDEVEYRKVNDAAASYKMLADSVRFKGYEKMEKSKKEQMKLEKMR ADFDQLELQKQIVDRAQSEIARLREBEENHHGGDDSEDEMEENDSDVNLSD" complement (join(3794..4276,4363..4490,4577..4673, 4757..5179,5354..5548,5693..5970,6088..6214,6297..6496, 6601..6648,6743..6776,6882..7013,7464..7559,7677..7743, 8162..8217,8290..8385,8470..8998,9122..9462)) /note="gene id:MOB24.2" /codon_start=1 /evidence=not_experimental /product="DNA_repair protein MutS" /protein_id="BAB01996.1" /db_xref="GI:9294039" /translation="MQQRSLTSFFQKFTATTGKLVSGDAAGGGGGGPRFNVKEG DAKGDASVRFAVSKSVDEVRGTDTPPEKVPVRVLPSPGPKPAESAGDASSLPNIMHKF VKVDDRDCSGERSREDVVLNDSSLCKANDVIFQFRSNNKTKQRNHAFFSFGRAEL
FEATURES source	1..79706 /organism="Arabidopsis thaliana" /mol_type="genomic DNA" /strain="Columbia" /db_xref="taxon:3702" /chromosome="3" /clone_lib="Mitsui P1" 2398..-3399 /note="gene id:MOB24.1 unknown protein" /codon_start=1 /evidence=not_experimental /protein_id="BAB01995.1" /db_xref="GI:9294038" /translation="MGSEDETGYPKFYSLNRONHPMYSRPIPKRHAYNEEDED VEGEEDPQGYIIGNERFQRKPKNVSGFEAGSDAKVADNRQEAFVLELV GDRILQRRSLRNEWNAEKVSEELRSEETQCRMTDDIKRKREKI KVEKS GLSSKWSFFNKDMLLCVSPKSLGACGVDSGEFVPMNTKVLVDKSNQDMDSP GDSEEDDEVEYRKVNDAAASYKMLADSVRFKGYEKMEKSKKEQMKLEKMR ADFDQLELQKQIVDRAQSEIARLREBEENHHGGDDSEDEMEENDSDVNLSD" complement (join(3794..4276,4363..4490,4577..4673, 4757..5179,5354..5548,5693..5970,6088..6214,6297..6496, 6601..6648,6743..6776,6882..7013,7464..7559,7677..7743, 8162..8217,8290..8385,8470..8998,9122..9462)) /note="gene id:MOB24.2" /codon_start=1 /evidence=not_experimental /product="DNA_repair protein MutS" /protein_id="BAB01996.1" /db_xref="GI:9294039" /translation="MQQRSLTSFFQKFTATTGKLVSGDAAGGGGGGPRFNVKEG DAKGDASVRFAVSKSVDEVRGTDTPPEKVPVRVLPSPGPKPAESAGDASSLPNIMHKF VKVDDRDCSGERSREDVVLNDSSLCKANDVIFQFRSNNKTKQRNHAFFSFGRAEL
CDS	10131..10577 /note="gene id:MOB24.3" /codon_start=1 /evidence=not_experimental /product="ethylene-responsive transcriptional coactivator-like protein" /protein_id="BAB01997.1" /db_xref="GI:9294040" /translation="MPSRYPGAVTDWEPVVLHKSQKODLRDPKAVNALRNGVAV QTVKFDAGSNKKGKSTAVFVINTKLEESTEPAMDVRVAEVRMLQKARLEKMSQ ADLAKQINERTQVQYGVKAVPNQVLAEMKRGVLGVKRGKIGK" complement (join(11108..11381,11467..11528,11699..11836, 11934..12071,12161..12334,12421..12510,12718..13104, 13623..13765,14306..14405)) /note="gene id:MOB24.4" /codon_start=1 /evidence=not_experimental /product="aldehyde dehydrogenase" /protein_id="BAB01998.1" /db_xref="GI:9294041" /translation="MENGKNGATTVKLPEIKETKLFINGOFIDASGKTFETIDPN GEVITAEGDEKEDVDLAVNAFYADHDPWPRTGFERAKLINKEFADLLEENIELA KLDAVDGKGLFQLGKYADIPATAGHFRYNAGAADKHGETLMTROSLFGYTLKEPIG VVGNIIPWNPFSIMFATKVPAMAAAGCTVWVPAEQTSLSALFYAHLKSEAGIPDGLV NIVTFGSGTAAAIASHMDVDKVSFTGSDVGRKIMQAAAANLKKVSLBELGKSPLL IFNDADIDKAADLALGCFYNGEICVASSRFVQEGIDYKVVVKLEKAKDWTVGDP FDSTARQGPQVDKROPEKILSYIEHGKNEGATLLTGGAIGDKGYFIQPTIEADVTED MKIYQDIEFPVMSLMKPKTVBEGIKCANNTKGLAAGILSQDIDLINTVSRSIKAGI IWNCYRFGDLDPCYGYKMSGNCRESGMDALDNIQTGSVVMPLHNSPMW" complement (join(19742..19933,20400..20657)) /note="gb AAD25142.1 gene id:MOB24.5 similar to unknown protein" /codon_start=1 /evidence=not_experimental /protein_id="BAB01999.1" /db_xref="GI:9294042" /translation="MTPLSISAGSPWPLLSPPLSSSLNRAAPSIFFPSFRINHGF SCSFIQVQALSRKTKQIQVVMAPPEEKLTRNLPDPIEWRPKGRRPDIFFK FSPKSPLPMPDYDPPEEEEEEKEETEDDPEKEDDQDNRS" join(23402..23459,23819..24021) /note="gene id:MOB24.6 unknown protein" /codon_start=1 /evidence=not_experimental /protein_id="BAB02000.1" /db_xref="GI:9294043" /translation="MKNASLKJLLFILVITSNLGAERKLTGVADVIVEGEAASPO YGIIDENLTKSPSKRDICTFQCRRGFCNLILQKCECL" complement (join(25945..26640,26714..26824,26914..27012, 27119..27224,27291..27360,27436..27583)) /note="contains similarity to ubiquitin conjugating enzyme gene id:MOB24.7" /codon_start=1 /evidence=not_experimental /protein_id="BAB02001.1"
CDS	10131..10577 /note="gene id:MOB24.3" /codon_start=1 /evidence=not_experimental /product="ethylene-responsive transcriptional coactivator-like protein" /protein_id="BAB01997.1" /db_xref="GI:9294040" /translation="MPSRYPGAVTDWEPVVLHKSQKODLRDPKAVNALRNGVAV QTVKFDAGSNKKGKSTAVFVINTKLEESTEPAMDVRVAEVRMLQKARLEKMSQ ADLAKQINERTQVQYGVKAVPNQVLAEMKRGVLGVKRGKIGK" complement (join(11108..11381,11467..11528,11699..11836, 11934..12071,12161..12334,12421..12510,12718..13104, 13623..13765,14306..14405)) /note="gene id:MOB24.4" /codon_start=1 /evidence=not_experimental /product="aldehyde dehydrogenase" /protein_id="BAB01998.1" /db_xref="GI:9294041" /translation="MENGKNGATTVKLPEIKETKLFINGOFIDASGKTFETIDPN GEVITAEGDEKEDVDLAVNAFYADHDPWPRTGFERAKLINKEFADLLEENIELA KLDAVDGKGLFQLGKYADIPATAGHFRYNAGAADKHGETLMTROSLFGYTLKEPIG VVGNIIPWNPFSIMFATKVPAMAAAGCTVWVPAEQTSLSALFYAHLKSEAGIPDGLV NIVTFGSGTAAAIASHMDVDKVSFTGSDVGRKIMQAAAANLKKVSLBELGKSPLL IFNDADIDKAADLALGCFYNGEICVASSRFVQEGIDYKVVVKLEKAKDWTVGDP FDSTARQGPQVDKROPEKILSYIEHGKNEGATLLTGGAIGDKGYFIQPTIEADVTED MKIYQDIEFPVMSLMKPKTVBEGIKCANNTKGLAAGILSQDIDLINTVSRSIKAGI IWNCYRFGDLDPCYGYKMSGNCRESGMDALDNIQTGSVVMPLHNSPMW" complement (join(19742..19933,20400..20657)) /note="gb AAD25142.1 gene id:MOB24.5 similar to unknown protein" /codon_start=1 /evidence=not_experimental /protein_id="BAB01999.1" /db_xref="GI:9294042" /translation="MTPLSISAGSPWPLLSPPLSSSLNRAAPSIFFPSFRINHGF SCSFIQVQALSRKTKQIQVVMAPPEEKLTRNLPDPIEWRPKGRRPDIFFK FSPKSPLPMPDYDPPEEEEEEKEETEDDPEKEDDQDNRS" join(23402..23459,23819..24021) /note="gene id:MOB24.6 unknown protein" /codon_start=1 /evidence=not_experimental /protein_id="BAB02000.1" /db_xref="GI:9294043" /translation="MKNASLKJLLFILVITSNLGAERKLTGVADVIVEGEAASPO YGIIDENLTKSPSKRDICTFQCRRGFCNLILQKCECL" complement (join(25945..26640,26714..26824,26914..27012, 27119..27224,27291..27360,27436..27583)) /note="contains similarity to ubiquitin conjugating enzyme gene id:MOB24.7" /codon_start=1 /evidence=not_experimental /protein_id="BAB02001.1"
CDS	10131..10577 /note="gene id:MOB24.3" /codon_start=1 /evidence=not_experimental /product="ethylene-responsive transcriptional coactivator-like protein" /protein_id="BAB01997.1" /db_xref="GI:9294040" /translation="MPSRYPGAVTDWEPVVLHKSQKODLRDPKAVNALRNGVAV QTVKFDAGSNKKGKSTAVFVINTKLEESTEPAMDVRVAEVRMLQKARLEKMSQ ADLAKQINERTQVQYGVKAVPNQVLAEMKRGVLGVKRGKIGK" complement (join(11108..11381,11467..11528,11699..11836, 11934..12071,12161..12334,12421..12510,12718..13104, 13623..13765,14306..14405)) /note="gene id:MOB24.4" /codon_start=1 /evidence=not_experimental /product="aldehyde dehydrogenase" /protein_id="BAB01998.1" /db_xref="GI:9294041" /translation="MENGKNGATTVKLPEIKETKLFINGOFIDASGKTFETIDPN GEVITAEGDEKEDVDLAVNAFYADHDPWPRTGFERAKLINKEFADLLEENIELA KLDAVDGKGLFQLGKYADIPATAGHFRYNAGAADKHGETLMTROSLFGYTLKEPIG VVGNIIPWNPFSIMFATKVPAMAAAGCTVWVPAEQTSLSALFYAHLKSEAGIPDGLV NIVTFGSGTAAAIASHMDVDKVSFTGSDVGRKIMQAAAANLKKVSLBELGKSPLL IFNDADIDKAADLALGCFYNGEICVASSRFVQEGIDYKVVVKLEKAKDWTVGDP FDSTARQGPQVDKROPEKILSYIEHGKNEGATLLTGGAIGDKGYFIQPTIEADVTED MKIYQDIEFPVMSLMKPKTVBEGIKCANNTKGLAAGILSQDIDLINTVSRSIKAGI IWNCYRFGDLDPCYGYKMSGNCRESGMDALDNIQTGSVVMPLHNSPMW" complement (join(19742..19933,20400..20657)) /note="gb AAD25142.1 gene id:MOB24.5 similar to unknown protein" /codon_start=1 /evidence=not_experimental /protein_id="BAB01999.1" /db_xref="GI:9294042" /translation="MTPLSISAGSPWPLLSPPLSSSLNRAAPSIFFPSFRINHGF SCSFIQVQALSRKTKQIQVVMAPPEEKLTRNLPDPIEWRPKGRRPDIFFK FSPKSPLPMPDYDPPEEEEEEKEETEDDPEKEDDQDNRS" join(23402..23459,23819..24021) /note="gene id:MOB24.6 unknown protein" /codon_start=1 /evidence=not_experimental /protein_id="BAB02000.1" /db_xref="GI:9294043" /translation="MKNASLKJLLFILVITSNLGAERKLTGVADVIVEGEAASPO YGIIDENLTKSPSKRDICTFQCRRGFCNLILQKCECL" complement (join(25945..26640,26714..26824,26914..27012, 27119..27224,27291..27360,27436..27583)) /note="contains similarity to ubiquitin conjugating enzyme gene id:MOB24.7" /codon_start=1 /evidence=not_experimental /protein_id="BAB02001.1"
CDS	10131..10577 /note="gene id:MOB24.3" /codon_start=1 /evidence=not_experimental /product="ethylene-responsive transcriptional coactivator-like protein" /protein_id="BAB01997.1" /db_xref="GI:9294040" /translation="MPSRYPGAVTDWEPVVLHKSQKODLRDPKAVNALRNGVAV QTVKFDAGSNKKGKSTAVFVINTKLEESTEPAMDVRVAEVRMLQKARLEKMSQ ADLAKQINERTQVQYGVKAVPNQVLAEMKRGVLGVKRGKIGK" complement (join(11108..11381,11467..11528,11699..11836, 11934..12071,12161..12334,12421..12510,12718..13104, 13623..13765,14306..14405)) /note="gene id:MOB24.4" /codon_start=1 /evidence=not_experimental /product="aldehyde dehydrogenase" /protein_id="BAB01998.1" /db_xref="GI:9294041" /translation="MENGKNGATTVKLPEIKETKLFINGOFIDASGKTFETIDPN GEVITAEGDEKEDVDLAVNAFYADHDPWPRTGFERAKLINKEFADLLEENIELA KLDAVDGKGLFQLGKYADIPATAGHFRYNAGAADKHGETLMTROSLFGYTLKEPIG VVGNIIPWNPFSIMFATKVPAMAAAGCTVWVPAEQTSLSALFYAHLKSEAGIPDGLV NIVTFGSGTAAAIASHMDVDKVSFTGSDVGRKIMQAAAANLKKVSLBELGKSPLL IFNDADIDKAADLALGCFYNGEICVASSRFVQEGIDYKVVVKLEKAKDWTVGDP FDSTARQGPQVDKROPEKILSYIEHGKNEGATLLTGGAIGDKGYFIQPTIEADVTED MKIYQDIEFPVMSLMKPKTVBEGIKCANNTKGLAAGILSQDIDLINTVSRSIKAGI IWNCYRFGDLDPCYGYKMSGNCRESGMDALDNIQTGSVVMPLHNSPMW" complement (join(19742..19933,20400..20657)) /note="gb AAD25142.1 gene id:MOB24.5 similar to unknown protein" /codon_start=1 /evidence=not_experimental /protein_id="BAB01999.1" /db_xref="GI:9294042" /translation="MTPLSISAGSPWPLLSPPLSSSLNRAAPSIFFPSFRINHGF SCSFIQVQALSRKTKQIQVVMAPPEEKLTRNLPDPIEWRPKGRRPDIFFK FSPKSPLPMPDYDPPEEEEEEKEETEDDPEKEDDQDNRS" join(23402..23459,23819..24021) /note="gene id:MOB24.6 unknown protein" /codon_start=1 /evidence=not_experimental /protein_id="BAB02000.1" /db_xref="GI:9294043" /translation="MKNASLKJLLFILVITSNLGAERKLTGVADVIVEGEAASPO YGIIDENLTKSPSKRDICTFQCRRGFCNLILQKCECL" complement (join(25945..26640,26714..26824,26914..27012, 27119..27224,27291..27360,27436..27583)) /note="contains similarity to ubiquitin conjugating enzyme gene id:MOB24.7" /codon_start=1 /evidence=not_experimental /protein_id="BAB02001.1"


```

/db_xref="GI:9294044"
/translation="MAQARLSRMQKELKLLSDPPHGASFPHLSSAANGSGDFSTF
STIDAQLEGPDVYANGIENLKIQIPERYPQPIVSEFATPIVHPNIDNSGRICLDI
LNLPPKANQWSPNLSITVLSMRLLLSSEPNPDDGLMCEVSEYKYNQTFDYKAREMT
EDNGKRAVGVGEGNSFGNDGKIRKSKLSTALPSQKQKDCGEBEILTRGVSAAC
KENKPNLNGKLSLGLKQPCNDTLASPTSAKSDNNRSLRKLSLRSPGLGELNEVSK
PEVLAQTDMLKEMQNDARSLRGEFENSVLEETSMABSIIVLSDSDSGQEBEERVS
SRSLSLAKRVLAKRP"

```

CDS

```

28636..30168
/note="emb|CAB79678.1
gene id:MOB24.8
similar to unknown protein"
/codon_start=1
/evidence=not experimental
/protein_id="BAB02002.1"
/db_xref="GI:9294045"

```

```

/translation="MSSLIAVLSFNSFPLKLSRILSNLMLLVRTSTWIKRKCWL
AIPYHLHATPLWKSITMLLRFSQMVPEPTTYFEISLSRAFILQAMHGFTLLWD
TTSFGLTRILWFSSAPHGQOQTSTFPSPENDASLSSPYNRRHHDLHLSSEGLK
GLQVLGLVFGPLAYRMPISKDDPKLPLKPRSDLYSLSTFSQSLVHLLASIDRHH
TNHNLVLIHLKNSLTPPLVSSRTFTVGFCLKNGIMEPTPTSRWCDCLPASATL
CISPTHGISQKSCISRVKAHLSRISKEVPNLFELITEPSPHALVVKGLKIGMI
TSPSGDYQRFNLHSHLYPLTEPKQRFNIIKSIHDLQSLGKNGIMISPSWDDYR
SFFNLDLPTIVTHLGDQQLLVLSARTSALEPYSTSRLLTVTNLLSIDSFED
PSTTRDLTCAIMPYSYMLKALNFMFSFHIIIVHILVGNAGFNVYVNFESLILIFV
LI"

```

CDS

```

join(32831..33058,33143..33907)
/note="contains similarity to heat shock transcription
factor"
gene id:MOB24.9"

```

```

/codon_start=1
/evidence=not experimental
/protein_id="BAB02003.1"
/db_xref="GI:9294046"

```

```

/translation="MEDNSNNNNNNVIAPFIVKTYQMVNDPSTDWLITWGPANHNSF
IVDPDLFSORILPFAKNNFNSFVQLANTYFRKVDPPDRWEPANFHLRQKQHLN
NIAARKHAGMYGDGEIVREIERLKEORELEABEIORNNRRIEATKRPEQMA
FLYKVVDPDLLPMLMEKERTKQOQVSDKKERVTMTSVKSEEREVEDEGRVFRV
MSSSTPSPSTENLYRHSFDGWLVPMTQOQFSGYETGLVAKSMLNSTSTSSSLTS
TFSLPESVGGGGGCGSIOGERRYKETATFGVVENPPTTPPYPSLFRGGF"
complement(join(37054..37590,37670..37809,37894..38197,
38288..38398,38486..38584,38662..38772,39019..39066,
39311..39367))

```

CDS

```

/note="gene id:MOB24.10
unknown protein"
/codon_start=1

```

Alignment Scores:

```

Pred. No.: 7,29e-55 Length: 79706
Score: 2622.00 Matches: 573
Percent Similarity: 61.19% Conservative: 23
Best Local Similarity: 58.83% Mismatches: 43
Query Match: 75.93% Indels: 337
DB: 8 Gaps: 19

```

US-10-086-464-2 (1-647) x AB020746 (1-79706)

```

Qy 1 MetSerAlaProSerProGlyThr-----GlySerProProSerProSer 17
Db 51787 ATGTCCAGCGCGTCTCCAGGCACTACTCCATCACCCTCCACCGTCTCTCCCA 51846
Qy 18 AsnSerThrThrProProAla-----SerAlaProProThrThrPro 35
Db 51847 AACTCGAACCAACCACTCTCTCCAGACGCTCTCTCTCTCCACCAACCTCT 51906
Qy 36 SerSerProProPro---ProSerThrIleProThrSerProProProSerArgSer 54
Db 51907 TCTTCTCTCTCGTCGCGCTCAACTAATTCACCTCTCTCTCTCTCTCTCTCT 51966
Qy 55 ThrProSerAlaProProProSerProProThrProThrProGlySerProProPro 74
Db 51967 CCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 52014

```

```

Qy 75 LeuProGlnProSerProProAlaProThrThrProGlySerProProAlaProValThr 94
Db 52015 CTTCCTCAACTTCCCTCCGCTCCGCTCCATCACTCTCT---TCTCCACCGTCTCCACCA 52071
Qy 95 ProProThr---ArgAsnProProProSerValProGlyProProSerAsnProSerArg 113
Db 52072 CCCTCAAAACCTCGAAGCCCTCCATCTCTTAACAAGGAGCACCAACACATCCCTCA 52128
Qy 114 GluGlyGlySerProArgProProSerSer-----ProSerProProProSerSer 131
Db 52129 ---GGATCTACTCTAGAACCTCAAACTAAACCGTCGCGCGCTGTGATCTCTCC 52185
Qy 132 AspGlyLeuSerThrGlyValValValGlyIleAlaIleGlyGlyValAlaLeuLeuVal 151
Db 52186 GATGGATTGTCACCGGAGTTGTGTAGGAATGCCCATTCGGAGGAGTCGTATTCTTGT 52245
Qy 152 IleValThrLeuLeuCysLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 169
Db 52246 ATACTGACTCTGATTTGCTCTCTCTGTAAGAAGAAAGAAAGAAAGACACGACGATGAA 52305
Qy 170 AspAlaTyrrTyrrValProProProProProGlyProGlyProGlyProGlyPro 184
Db 52306 GCTGCTTACTATGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 52364
Qy 184 ----- 184
Db 52365 CAATTCATGAATAAATAATCGAATATTTAAACCATAGTTGTTCAATTCGTATACGATTG 52424
Qy 184 ----- 184
Db 52425 CTCAAATCTGATCTACTGATTATGCAATTGTTATCTGTTAAAGAGATTACCTTTGTGGA 52484
Qy 185 -----GlyGlyProTyrrGlyGlyGln 191
Db 52485 GTTTCAAACTTTACAGAGATTGATGTTGTTTATCAGCTCGAGGACCTTACGGTGGTCAA 52544
Qy 192 GlnGlnGlnTrpArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 210
Db 52545 CAACAGTATTGGCAACACACAAACCGCTCAGCGCGTCAGATATCATGTAGTAGCTCA 52604
Qy 211 LeuProProProProProAlaProSerProProArgGlnProProProProProPro 230
Db 52605 TTG---CCACCACCTAAGCCTCATCTCCACCAAGAAACCTCTCTCGCCACCTCCACCA 52661
Qy 231 Pro---PheMetSerSerSerGlyGlySerAspTyrrSerAspArgProValLeuPro 249
Db 52662 CCAGCATTCATAGTAGTACAGTGTGTGTTCTGACTATTTCGGATCTTCCGGTCTTCTCCA 52721
Qy 250 ProSerProGlyLeuValLeuLeuGlyPheSerLysSerThrPheThrTyrrGluGluLeuAla 269
Db 52722 CCATCTCCAGGCGTTGTGTAGGCTTTTCTAAAGACACTTTCACTATGAGGAGTTGTCG 52781
Qy 270 ArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyGlyPheGlyTyrrVal 289
Db 52782 AGAGCTACTAATGCTTCTCTGAGGCTAAATTTGTAGCAAGAGGAGGTTGTTATGTG 52841
Qy 290 HisLysGlyValLeuProSerGlyLysGluValAlaValLysGlnLeuLysValGlySer 309
Db 52842 CATAAAGGTATATGTGCTAGTGGGAAAGAAAGTTGCTGTGAAACAGATTGAAAGCTGGTAGT 52901
Qy 310 GlyGlnGlyGluArgGluPheGlnAlaGluValGluIleLeuSerArgValHisHisArg 329
Db 52902 GGTACGGAGAGAGAGAGTTTCAGGCTGAGTTGAGTATCATTAGCAGAGTTTCATCAGG 52961
Qy 330 HisLeuValSerLeuValGlyTyrrCysIleAlaGlyAlaLysArgLeuLeuValTyrrGlu 349
Db 52962 CATTTGGTTTCTCTTATTGTTGTTTATTGTTGCGCGTGTCAAAGATTACTTGTCTATGAG 53021
Qy 350 PheValProAsnAsnAsnLeuLeuLeuHisLeuHisGlyGlu----- 363
Db 53022 TTTTGTCCAAACCAACATCTTGAGTTTCACCTCCATCGTTTA-GAACAAAACTAAACTCTT 53080

```

QY	363	-----	363
Db	53081	TCCTTTGATTTTGTACATGCTTTGGTTGTTTTCAGTTAGTCGATGATTGTTGGGAATT	53140
QY	364	-----GlyArgProThrMetGluTrpSerThrArgLeuLysIleAla	377
Db	53141	GAATTTGTTATAGTTAAGGACGGCTACGATGGATGGATGACTAGATTTGAAGATTGCT	53200
QY	378	LeuGlySerAlaLysGlyLeuSerTyrLeuHisGluAspCys-----	391
Db	53201	CTTGATCTGCTAAAGGACTTTTCATATCTTCATGAAGATTGTAAGTCTFAAAACCTTAAA	53260
QY	391	-----	391
Db	53261	ATTATGTACAGATGTTTGATAAATACTGCTTTTGTGTTGTCACAGGATATATATAAAT	53320
QY	392	-----AnProLysIleIleHisArgAspIleLysAlaSerAsnIleLeuIleAs	408
Db	53321	ATTTCGAGCGAATCCGAANAATCAATCCAGTGATATTAAAGCGTCAAAACATATTGATTGA	53380
QY	408	pPheLysPheGluAlaLys-----	414
Db	53381	TTTCAAATTTGAAGCTAAGGTATGTTTACTCGATATGTTTGTGAAGTGTGTTTGTAAAT	53440
QY	415	-----ValAlaAspPheGlyLeuAlaLysIleAlaAs	425
Db	53441	TGTGTCACCTAAAGCTCTTTCTTTTCAGGTTGCTGACTTTTGGTCTTGCCGAAGATTGCTT	53500
QY	425	erAspThrAsnThrHisValSerThrArgValMetGlyThrPheGly-----	440
Db	53501	CTGATACAAACACTCATCTATGTCACACGCGTGATGGGAACCTTTGGGTAAANAATCATCTT	53560
QY	440	-----	440
Db	53561	TAATCTACTAAACATGTACATTTATGTCAGTTGTTGTTTAATGTCATGTCATCTTGTGCAACA	53620
QY	441	-----TyrLeuAlaProGluTyrAlaAlaSerG	450
Db	53621	AGTTCCTGAAACTTTTGGTTTGGGATTTTAGGTATTTGCGTCCGGAATATCTGCAAGTG	53680
QY	450	lyLysLeuThrGluLysSerAspValPheSerPheGlyValValLeuLeuGluLeuIleT	470
Db	53681	GAAGCTCACAGAAAAGTCTGACGTTTCTCATTTGGCGTTGACTTTTGGACTTATTA	53740
QY	470	hrGlyArgArgProValAspAlaAsnValTyrValAspAspSerLeuValAspTrp-	489
Db	53741	CTGGAGGCGCCTGTTGATGCGAACATGCTATGTAGATGACAGCTTAGTTGACTGGG	53800
QY	489	-----	489
Db	53801	TACTTTACTATGACTGCAGATAAATCTAAACTTCTGATTTTCTACAAATTTAACTGTTGAC	53860
QY	490	-----AlaArgProLeuLeuAsnArgAla	497
Db	53861	TAAACAAAAGCTATGTTGATGTTGATGTTGGAGACGGCACCACTTCTTTAACCGAGCA	53920
QY	498	SerGluGlnGlyAspPheGluGlyLeuAlaAspAlaLysMetAsnAsnGlyTyrAspArg	517
Db	53921	TCTGAGGAAGGAGATTTTGGGGTTTGGCTGATTCAAAGATGGGTAAATGAGTATGACAGA	53980
QY	518	GluGluMetAlaArgMetValAlaCysAlaAlaAlaCysValArgHisSerAlaArgArg	537
Db	53981	GAGGAGATGGCTCGCATGTTGCTTTCGCTCGCGCTGCTGTTGTCGCCATTCAGCTCGCGGC	54040
QY	538	ArgProArgMetSer-----	542
Db	54041	AGACCTCGCATGAGCCAGGTAACTTAATCCTTTCTCTAGAGAGAGACATATTGTTGG	54100
QY	542	-----	542
Db	54101	TTCTGTTTATTGGTGAATAATAATGTTTCCGTATGTTTAAAGAAAGGTTTGTGATG	54160
QY	542	-----	542
Db	54161	GTTTTTGTGAGAAAGATCGGTTTCAGTTATGTTCTTATGCTACGCCCTACTCTTTTGTACC	54220
QY	542	-----	542
Db	54221	CATCATCATATCTATAACAACAACACCGATTGAATATGTCAAAAGAAATTTGGCTGCAGTTTC	54280
QY	542	-----	542
Db	54281	TTTTGTTTTCATTTTGTCCAAAGGATTGTTCATTTTAAACATATGTTTAATGCTAATGGAGTGAG	54340
QY	543	-----GlnIleValArgAlaLeuGluGlyAsnValSerLeuSer	555
Db	54341	TGATTTCTGCTACAAAATAGCAGATAGTAGCGGCGTTAGAGGAAATGATATCTGCTGCT	54400
QY	556	AspLeuAsnGluGlyMetArgProGlyGlnSerAsnValTyrSerSerTyrGlyGlySer	575
Db	54401	GATCTTAAACGAAGGATGAGACCGGGTCACAGCAACGTATACAGTCTATATGGAGGAAGC	54460
QY	576	ThrAspTyrAspSerSerGlnTyrAsnGluAspMetLysPheArgLysMetAlaLeu	595
Db	54461	ACAGACTATGACAGCGCAATACACAGCAGCATGATAAAGTTTAGGAAAATGGCTCTT	54520
QY	596	GlyThrGlnGluTyrAsnAlaThrGlyGluTyrSerAsnProThrSerAspTyrGlyLeu	615
Db	54521	GGAACTCAAGAAATACGCCCAACCGCGGAGTACAGTAAATCAACAGTACTACGAGCTG	54580
QY	616	TyrProSerGlySerSerSerGluGlyGlnThrThrArgGluMetGluMetGlyLysIle	635
Db	54581	TACCCGCTCGTTCAACGAGTGAAGGTCAAGCCACACGAGAAATGGAGATGGGAAAGATT	54640
QY	636	LysArgThrGlyGlnGlyTyrSerGlyProSerLeu	647
Db	54641	AAGAAAACCGTCAAGGTTTATAGTGGACCTCTCTT	54676
RESULT 15			
LOCUS	AK103247		
DEFINITION	Oryza sativa (japonica cultivar-group) cDNA clone:J033123K23, full insert sequence.	2307 bp	linear
ACCESSION	AK103247		
VERSION	AK103247.1	GI:32988456	
KEYWORDS	FLU_CDNA; CAP trapper.		
SOURCE	Oryza sativa (japonica cultivar-group)		
ORGANISM	Oryza sativa (japonica cultivar-group)		
REFERENCE	1		
AUTHORS	The Rice Full-length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-length cDNA Project Team, Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., Foundation of Advancement of International Genome Sequencing & Analysis Group, Otono, Y., Murakami, K., Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y., Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M., Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J., Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J., Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN, Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Oeato, N., Oka, Y., Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Yoshino, M. and Hayashizaki, Y.		
TITLE	Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice		
JOURNAL	Science 301 (5631), 376-379		(2003)
MEDLINE	22752273		
PUBMED	12869764		
REFERENCE	2		(bases 1 to 2307)
AUTHORS	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K.,		

Job time : 5988 secs

```
Db      988  |||||TATGTTCAAAAGAGNTTCTCCGATCGGACAGAGGTTGCTGTGAACCAATTGAGAGAT 1047
Qy      308  GlySerGlyGlnGlyArgGluPheGlnAlaGluValGluIleIleSerArgValHis 327
Db     1048  GGGAGTGGCAGGAGGAGCGTGAATTTTCAGCGGAGGTTGAGATTATCAGCGGGTACAT 1107
Qy     328  HisArgHisLeuValSerLeuValGlyTyrCysIleAlaGlyAlaLysArgLeuVal 347
Db     1108  CACAAGCATCTTGTAAACATTGGTGTATTGTCATTTCTGGAGGGAAGAGGTTGCTTGTG 1167
Qy     348  TyrGluPheValProAsnAsnLeuGluLeuHisLeuHisGlyGluGlyArgProThr 367
Db     1168  TATGATGATGTTCCAAACACACATTGGAGCTCCACTTCATGGAGAGCGCCGACCAACC 1227
Qy     368  MetGluTrpSerThrArgLeuLysIleAlaLeuGlySerAlaLysGlyLeuSerTyrIleu 387
Db     1228  ATGGAATGGCTTACAAGATTACGCATTGCTTTGGTGTCTGCGAAGGGATTGGCATATCTT 1287
Qy     388  HisGluAspCysAsnProLysIleIleHisArgAspIleLysAlaSerAsnIleIle 407
Db     1288  CATGAAGATTGCCATCTCAAGATCATTCCCGTGACATATAAAGTCAGCAAAATATTCTTCT 1347
Qy     408  AspPheLysPheGluAlaLysValAlaAspPheGlyLeuAlaLysIleAlaSerAspThr 427
Db     1348  GACCGGAGATTGAGGCAAGGTGGCAGATTTCCGACTTGCATAAATTGACTTCTGATAAT 1407
Qy     428  AsnThrHisValSerThrArgValMetGlyThrPheGlyTyrLeuAlaProGluTyrAla 447
Db     1408  AACACTCATGTTTCAACCAGAGTAATGGGCACATTTGGGTACCTTGCACACAGAGTATGCG 1467
Qy     448  AlaSerGlyLysLeuThrGluLysSerAspValPheSerPheGlyValValLeuLeuGlu 467
Db     1468  TCTTCTGGCCAGCTAACTAGAGAAATCAGATGCTCTTTCTTCGGAGTAATGCTTCTTGAG 1527
Qy     468  LeuIleThrGlyArgProValAspAlaAsnAsnValTyrValAspAspSerLeuVal 487
Db     1528  CTAATAACTGGCCGCGCTCTGTGAGATCGAACCATCACAGATGGATGACAGCTTGTT 1587
Qy     488  AspTrpAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGluGlyLeuAla 507
Db     1588  GATTGGCAAGGCCCTTTAATGATGCGAGCATCAGATGATGCAATTATGATGCTTTAGTG 1647
Qy     508  AspAlaLysMetAsnAsnGlyTyrAspArgGluGluMetAlaArgMetValAlaCysAla 527
Db     1648  GATCCACGGCTAGGACAGAGTACAATGGTTAACGAGATGGCGAGGATGATTGCTTGTGCA 1707
Qy     528  AlaAlaCysValArgHisSerAlaArgArgProArgMetSerGlnIleValArgAla 547
Db     1708  GCTGCTTGTTGACGCCATTCTGCACCGCAGCGCCACCGATGAGCCAGGTTGTTGGGCC 1767
Qy     548  LeuGluGlyAsnValSerLeuSerAspLeuAsnGluGlyMetArgProGlyGlnSerAsn 567
Db     1768  TTGGAAGCGCATGTGTCATTGTCATCTGAACGAAGGTTTCTGCTCTGGCCATAGCGGA 1827
Qy     568  ValTyrSerSerTyrGlySerThrAspTyrAspSerSerGlnTyrAsnGluAspMet 587
Db     1828  TTCTTTGGGATCATAC---AACAGCAACGAGTACGACACGGGCCATTACAATGAAGACCTG 1884
Qy     588  LysLysPheArgLysMetAlaLeuGlyThrGlnGluTyrAsnAlaThrGlyGluTyrSer 607
Db     1885  AAAAAATTCAGGAAGATGGCATTGGCAGTGGCAATCTA-----GAGAGCAGC 1932
Qy     608  AsnProThrSerAspTyrGlyLeuTyrProSerGlySerSerSerGluGlyGlnThrThr 627
Db     1933  CAGCAACCGCAGCCACCGAGTTCGTCCCGAATCGATCCGTATCAATGGGGGATCGCGGT 1992
Qy     628  Arg-----GluMetGluMetGlyLysIleLysArgThrGly 639
Db     1993  CAAATACCGGAGACGGAGATGGGAGTTTGAAGAAAGATGGC 2040
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 24, 2004, 18:28:17 ; Search time 604 Seconds

(without alignments)
4550.636 Million cell updates/sec

Title: US-10-086-464-2

Perfect score: 3453

Sequence: 1 MSSAPSGTSGPPSPSNT.....REMEMGXKRTGCGYSGPSL 647

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 212409041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum Match 100%

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10086464/runat_23042004_083014_5603/app.query.fasta_1.839
-DB=N Geneseq 29Jan04 -QMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10086464 -CGN 1 1 708 -runat_23042004_083014_5603 -NCPU=6 -ICPU=3
-NO_MMAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N Geneseq 29Jan04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002s:*
7: Geneseqn2003as:*
8: Geneseqn2003bs:*
9: Geneseqn2003cs:*
10: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query No.	Score	Match	Length	DB ID	Description
1	3453	100.0	1944	4	AAF77094	Asf77094 Brassica
2	3453	100.0	1944	9	ACF36548	B. napus
3	3453	100.0	2189	9	ACF36549	B. napus
4	2936.5	85.0	1959	9	ACF38557	ACF38557 A. thalia
5	2885	83.6	2324	3	AAC38968	Arabidops
6	1821.5	52.8	1812	9	ACF36554	O. sativa
7	1668	48.3	1515	9	ACF36555	Z. mays
8	1637.5	47.4	1488	9	ACF36556	G. max

9	1507	43.6	2520	4	AAF77096	Arabidops
10	1504	43.6	2520	9	ACF36551	Arabidops
11	1479.5	42.8	2820	4	AAF77097	Arabidops
12	1471.5	42.6	2820	9	ACF36552	Arabidops
13	1209.5	35.0	2880	4	AAF77095	Arabidops
14	1209.5	35.0	2880	9	ACF36550	Arabidops
15	1185	34.3	3060	9	ACF36553	Arabidops
16	1181	34.2	3060	4	AAF77098	Arabidops
17	839.5	24.3	1923	7	ADA70890	Rice gene
18	816.5	23.6	2124	6	ABZ13274	Arabidops
19	804.5	23.3	2865	7	ADA70701	Rice gene
20	802.5	23.2	2011	4	AAH22557	P. patens
21	797	23.1	1730	7	ADA70647	Rice gene
22	793.5	23.0	2823	7	ADA69478	Rice gene
23	782.5	22.7	1419	3	AAC43029	Arabidops
24	782.5	22.7	1419	6	ABZ13121	Arabidops
25	775.5	22.5	1626	3	AAC48313	Arabidops
26	775.5	22.5	1917	7	ADA70527	Rice gene
27	773	22.4	1479	3	AAC43040	Arabidops
28	769.5	22.3	1353	3	AAC42860	Arabidops
29	769.5	22.3	1353	6	ABZ14106	Arabidops
30	766	22.2	1494	3	AAC48560	Arabidops
31	758.5	22.0	1572	3	AAC42898	Arabidops
32	758.5	22.0	1572	6	ABZ13705	Arabidops
33	758.5	22.0	1920	6	AAL45905	A thalian
34	758.5	22.0	2012	3	AAC44847	Arabidops
35	756	21.9	3239	3	AA79271	Eucalyptu
36	754.5	21.9	2136	7	ADA69829	Rice gene
37	749.5	21.7	1751	3	AAC45313	Arabidops
38	744.5	21.6	1485	3	AAC42702	Arabidops
39	740.5	21.4	3525	7	ADA71170	Rice gene
40	740	21.4	2868	3	AA79266	Pinus rad
41	740	21.4	3153	7	ADA71066	Rice gene
42	736.5	21.3	1703	3	AAC46629	Zea mays
43	736	21.3	3624	7	ADA69512	Rice gene
44	735	21.3	2004	3	AAC40780	Arabidops
45	734	21.3	1455	3	AAC45877	Arabidops

ALIGNMENTS

RESULT 1
AAF77094
ID AAF77094 standard; DNA; 1944 BP.
AC AAF77094;
XX
XX 17-MAY-2001 (first entry)
DT
XX
XX Brassica napus PERK1 DNA.
XX
XX Proline-rich extensin-like receptor kinase; PERK; resistance; plant; ds.
OS Brassica napus.
XX
XX WO200114563-A1.
PN
XX
XX 01-MAR-2001.
PD
XX
XX 18-AUG-2000; 2000WO-CA000966.
PF
XX
XX 19-AUG-1999; 99US-0149466P.
PR
XX
XX 13-OCT-1999; 99US-0159122P.
XX
XX (GORI// GORING D.
PA (SILV// SILVA N.
PA
XX
XX Goring D, Silva N;
PI
XX
XX WPI; 2001-244305/25.
DR
XX
XX New proline-rich, extensin-like receptor kinase nucleic acids and
PT polypeptides useful for increasing plant wounding or pathogen resistance,

PT or for producing transgenic plants with increased wounding or pathogen
 PT resistance.

XX Claim 6; Fig 1; 91pp; English.

XX The present invention relates to proline-rich extensin-like receptor
 CC kinase (PERK). The PERK nucleic acids and polypeptides are useful for
 CC increasing the resistance of plants to wounding and pathogens. These are
 CC also useful for producing transgenic plants with increased wounding and
 CC pathogen resistance compared with a wild type plant, as well as in assays
 CC for identifying and developing compounds to inhibit and/or enhance
 CC polypeptide function directly

XX Sequence 1944 BP; 464 A; 550 C; 459 G; 471 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 4,62e-93 Length: 1944
 Score: 3453.00 Matches: 647
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 4 Gaps: 0

US-10-086-464-2 (1-647) x AAF77094 (1-1944)

Qy 1 MetSerSerAlaProSerProGlyThrGlySerProProSerProProSerAsnSerThr 20
 Db 1 ATGTCCTCGGGCGGCTCTCCGGGACTGGTTCCGCTCCATCTCCACCATCAAACTCCACA 60
 Qy 21 ThrThrThrProProAlaSerAlaProProProProThrProSerSerProProPro 40
 Db 61 ACCACCATCTCTCTCCAGCTTCGCGCTCTCTCCACCATCACTCTCTCTCCGCGG 120
 Qy 41 ProSerThrIleProThrSerProProProSerSerArgSerThrProSerAlaProPro 60
 Db 121 CCATCCACTATTCGACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180
 Qy 61 ProSerProProThrProSerThrProGlySerProProProProLeuProGlnProSerPro 80
 Db 181 CCATCTCCACCAACTCCATCTACGCGGGATCTCCACCTCTCTCTCTCTCTCTCTCTCCA 240
 Qy 81 ProAlaProThrThrProGlySerProProAlaProValThrProProThrArgAsnPro 100
 Db 241 CCCGCTTCAACTACGCGCGGATCTCCACCGGACCTGTCTCTCTCTCTCTCTCTCTCTCT 300
 Qy 101 ProProSerValProGlyProProSerAsnProSerArgGlyGlySerProArgPro 120
 Db 301 CCACCTTCAGTCCGAGACACCGTCCAACTCTTCACCGGAGGAGATCTCTCGACCT 360
 Qy 121 ProSerSerProProSerProSerSerSerSerSerSerSerSerThrGlyValValVal 140
 Db 361 CCATCTCTCTCGCGCGCT 420
 Qy 141 GlyIleAlaIleGlyGlyValAlaLeuLeuValIleValThrLeuIleCysLeuLeuCys 160
 Db 421 GGAATCGCCATCGGAGGAGTCT 480
 Qy 161 LysLysLysArgArgArgAspGluAspAlaTyrTyrValProProProProProProPro 180
 Db 481 AAGAAGAAACGACGAGACGAG 540
 Qy 181 GlyProLysAlaGlyGlyProTyrGlyGlyGlnGlnGlnGlnThrArgGlnGlnGlnAla 200
 Db 541 GGTCCCAAGCGGAGGAGCTTACCGTGGACAGCAGCAACAAATGGCGGCAACAAACGCA 600
 Qy 201 ThrProProSerAsnHisValThrSerLeuProProProProProProProProProPro 220
 Db 601 ACACACCGCTCAGATCATGTGTGACGTCACTTACCACCAACCACTTAAGGCTTCCATCTCCA 660
 Qy 221 ProArgGlnProProProProProProProProProProProProProProProProPro 240
 Db 661 CCACGCAACCTCTCTCCACCTCCACCGGCTTTTCATGACGACGAGCGGCGGCTCCGAC 720

Qy 241 TyrSerAspArgProValLeuProProProSerProGlyLeuValLeuGlyPheSerIys 260
 Db 721 TACTCGGACCGTCAGTTCTTCTCCACCGTCTCCAGGGCTTGTGTAGGCTTCTCCAAA 780
 Qy 261 SerThrPheThrTyrGluGluLeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeu 280
 Db 781 AGCACTTTCATACAGAGAGAGCTAGCTAGAGCCACCAATGGTTCTCCGAGCGCACTTG 840
 Qy 281 LeuGlyGlnGlyGlyPheGlyTyrValHisLysGlyValLeuProSerGlyLysGluVal 300
 Db 841 TTAGGACAAGCGGGTTCGGTTACGTGCACAAAGGTGTGTTCCTAGTGGGAAAGAGTT 900
 Qy 301 AlaValLysGlnLeuLysValGlySerGlyGlnGlyGluArgGluPheGlnAlaGluVal 320
 Db 901 GCTGTGAAGCAGTTGAAGTTGGGAGTGGTCAGGAGAGAGGGAGTTTCAGGCGAGAGGTT 960
 Qy 321 GluIleIleSerArgValHisHisArgHisLeuValSerLeuValGlyTyrCysIleAla 340
 Db 961 GAGATCATCAGCAGAGTTCCACACAGGCATCTGGTGTCTCTTGTGTGGTTATTGTCATCGCC 1020
 Qy 341 GlyAlaLysArgLeuLeuValTyrGluPheValProAsnAsnAsnLeuGluLeuHisLeu 360
 Db 1021 GGTGCCAAAAGATTGCTTGTCTATGAGTTTGTCTTAACAACAATCTCGAGCTTCACCTC 1080
 Qy 361 HisGlyGluGlyArgProThrMetGluTyrSerThrArgLeuLysIleAlaLeuGlySer 380
 Db 1081 CATGGCAGGAGCGGCTCAATGGAATGGAGCACCAGATTGAAGATTGCTCTTGGATCT 1140
 Qy 381 AlalysGlyLeuSerTyrLeuHisGluAspCysAsnProLysIleIleHisArgAspIle 400
 Db 1141 GCTAAGAGACTTCTTATCTTCATGAAGATTGCAATCTTAATCAATCAATCCCGTATATC 1200
 Qy 401 LysAlaSerAsnIleLeuIleAspPheGlyPheGluAlaLysValAlaAspPheGlyLeu 420
 Db 1201 AAGGCTTCAACAATATTGATAGATTTCAGTTTGAAGCTTAAGGTTGCTGATTTGGTCTT 1260
 Qy 421 AlalysIleAlaSerAspThrAsnThrHisValSerThrArgValMetGlyThrPheGly 440
 Db 1261 GCTAAGATTGCTCTGTATACAAACACGATGTATCAACAGCTGTGTGAGGAACTTTGGG 1320
 Qy 441 TyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPheSer 460
 Db 1321 TACTTGGCTCCGANTACGCTCAAGCGGAAAGCTCAGCGAGAGTCTGAGCTTTTCTCA 1380
 Qy 461 PheGlyValValLeuLeuGluLeuIleThrGlyArgArgProValAspAlaAsnVal 480
 Db 1381 TTTGGCGTTGTGCTTTTGGAGCTCATTTACTGAGCTCGACCGCTGATGCCCAACAATGTC 1440
 Qy 481 TyrValAspAspSerLeuValAspTyrAlaArgProLeuLeuAsnArgAlaSerGluGln 500
 Db 1441 TATGTATGACAGCTTAGTTGACTGGGCGACGACATTTGCTTAACCGAGCATCTGAGCAA 1500
 Qy 501 GlyAspPheGlyGlyLeuAlaAspAlaLysMetAsnAsnGlyTyrAspArgGluGluMet 520
 Db 1501 GGAGACTTTGAGGGTTTAGCTGATGCAAGATGAATTAATGGTATGACAGAGAGAGATG 1560
 Qy 521 AlaArgMetValAlaCysAlaAlaAlaCysValArgHisSerAlaArgArgProArg 540
 Db 1561 GCTCGATGGTGTGCTGTGCTGCGGCTTGTGTTCGCCATTCAGCTCGCGCAGACCTCGC 1620
 Qy 541 MetSerGlnIleValArgAlaLeuGluGlyAsnValSerLeuSerAspLeuAsnGluGly 560
 Db 1621 ATGAGCCAGATTGTGCGTGGCTTAGAAGGAAATGTATCATTGTCTCAGATCTTTAAACGAAGG 1680
 Qy 561 MetArgProGlyGlnSerAsnValTyrSerTyrGlyGlySerThrAspTyrAspSer 580
 Db 1681 ATGAGACCGAGTCAAGCAATGTATACAGCTCATACGAGGAGAACACCGATTATGACTCG 1740
 Qy 581 SerGlnTyrAsnGluAspMetLysLysPheArgLysMetAlaLeuGlyThrGlnGluTyr 600
 Db 1741 AGCCAGTACAATGAAGACATGAAGAGTTTAGGAAAAATGGCACTTGGAACTCAAGAGTAC 1800
 Qy 601 AsnAlaThrGlyGluTyrSerAsnProThrSerAspTyrGlyLeuTyrProSerGlySer 620

```
Db 1801 AACGCCAGGGTACAGTAATCCGACCACTGACTGACTGACTGCTGGTTCA 1860
Qy 621 SerSerGluGlyGlnThrArgGluMetGluMetGlyLysIleLysArgThrGlyGln 640
Db 1861 AGCAGCGAGGGCCAAACACACAGCAATGGAGATGGGGAAGATTAAAGAACCGGTGAG 1920
Qy 641 GlyTyrSerGlyProSerIeu 647
Db 1921 GGTATAGTGGACCTTCTT 1941

RESULT 2
ACF36548
ID ACF36548 standard; cDNA; 1944 BP.
XX
AC ACF36548;
XX
DT 18-DEC-2003 (first entry)
XX
DE B. napus PERK1 receptor kinase encoding cDNA.
XX
KW PERK1; transgenic; plant; proline-rich extensin-like receptor kinase;
KW wound; pathogen resistance; plant growth; seed production; gene; ss.
XX
OS Brassica napus.
XX
FH Key Location/Qualifiers
FT CDS 1..1944
FT /tag= a
FT /trans_except= (pos: 1332..1334, aa: Phe)
FT /product= "PERK1"
FT /note= "Proline-rich Extensin-like Receptor Kinase"
XX
PN W02003072763-A1.
XX
PD 04-SEP-2003.
XX
PF 28-FEB-2003; 2003WO-CA000274.
XX
PR 28-FEB-2002; 2002CA-02373903.
PR 28-FEB-2002; 2002US-00086464.
XX
PA (GORI/) GORING D.
PA (SILV/) SILVA N.
PA (HAFF/) HAFFANI Y Z.
XX
PI Goring D, Silva N, Haffani YZ;
XX
WPI; 2003-712727/67.
DR P-PSDB; ABR82937.
XX
PT Producing a transgenic plant having an increased plant resistance, plant
PT growth or seed production comprises transforming a plant with a nucleic
PT acid molecule having a Proline-rich Extensin-like Receptor Kinase
PT activity.
XX
PS Claim 3; Fig 1A; 123pp; English.
XX
CC The invention relates to producing a transgenic plant having increased
CC plant height, number of branches, number of seed pods and/or seed
CC production compared to a non-transgenic plant, and/or quicker flowering
CC or later senescence compared to a non-transgenic plant. The method
CC involves transforming a plant with a vector including a Proline-rich
CC Extensin-like Receptor Kinase (PERK) nucleic acid molecule or a nucleic
CC acid molecule having PERK activity. The method, as well as the PERK
CC nucleic acid molecule and polypeptide, are useful in increasing plant
CC resistance to wounding and pathogens and in increasing plant growth and
CC seed production. The nucleic acid molecule and polypeptide may also be
CC used in producing transgenic plants or transgenic host cells. The present
CC sequence represents a cDNA encoding a B. napus PERK1 receptor kinase
CC polypeptide
XX
SQ Sequence 1944 BP; 464 A; 550 C; 459 G; 471 T; 0 U; 0 Other;
```

```
Alignment Scores:
Pred. No.: 4.62e-93 Length: 1944
Score: 3453.00 Matches: 647
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-10-086-464-2 (1-647) x ACF36548 (1-1944)

Qy 1 MetSerSerAlaProSerProGlyThrGlySerProProSerProProSerAsnSerThr 20
Db 1 ATGTCTCTCGCGCGCTCTCGGGGACTGGTTCGCTCCATCTCCACCATCAATCCACA 60
Qy 21 ThrThrThrProProAlaSerAlaProProThrThrProSerSerProProPro 40
Db 61 ACCACCACCTCTCTCTCAGCTTCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 120
Qy 41 ProSerThrThrProProSerProProSerSerArgSerThrProSerAlaProPro 60
Db 121 CCATCCACTATTCGACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180
Qy 61 ProSerProThrProSerThrProGlySerProProProLeuProGlnProSerPro 80
Db 181 CCATCTCCACCACTCCATCTAGCGCGGATCTCCACCTCTCTCTCTCTCTCTCTCT 240
Qy 81 ProAlaProThrThrProGlySerProProAlaProValThrProThrArgAsnPro 100
Db 241 CCGCTCCAACTACGCCCGGATCTCCACCGCGACTCTTACTCTCTCTCTCTCTCTCT 300
Qy 101 ProProSerValProGlyProProSerAsnProSerArgGluGlyGlySerProArgPro 120
Db 301 CCACCTTCAGTCCAGGACCACTCCATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 360
Qy 121 ProSerSerProSerProProSerProSerSerAspGlyLeuSerThrGlyValVal 140
Db 361 CCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 420
Qy 141 GlyIleAlaIleGlyGlyValAlaLeuLeuValIleValThrLeuIleCysLeuLeu 160
Db 421 GGAATCGCCATCGGAGGAGTCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 480
Qy 161 LysLysLysArgArgArgAspGluGluAspAlaTyrTyrValProProProProPro 180
Db 481 AAGAGAAACGACGAGAGAGAGAGAGATGCTTACTATGTTCCTCCGCCACCTCTCT 540
Qy 181 GlyProLysAlaGlyGlyProTyrGlyGlyGlnGlnGlnGlnTrpArgGlnGlnAsnAla 200
Db 541 GGTCCCAAGCCGAGGAGACCTTACGGTGGAGCAGCAACAATGGCGGCAACAAACGCA 600
Qy 201 ThrProProSerAspHisValValThrSerIeuProProProProProProProPro 220
Db 601 ACACCAACCGTCAGATCATGTGTCGACGTCACTACCACCACTTAAAGGTCTCTCCA 660
Qy 221 ProArgGlnProProProProProProPheMetSerSerSerGlyGlySerAsp 240
Db 661 CCACGGCAACTCTCTCCACTCTCCACCCGCTTTCATGACGACGACGGGGCTCCGAC 720
Qy 241 TyrSerAspArgProValLeuProProProSerProGlyLeuValLeuGlyPheSerLys 260
Db 721 TACTCGGACCGTCCAGTTCCTCTCTCCACCGCTCTCCAGGGCTTGTGTAGGCTTCT 780
Qy 261 SerThrPheThrTyrGluGluLeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeu 280
Db 781 AGCACTTTTCATACGAGGAGTAGCTAGAGCCACCAATGGTTTCTCCGAGGGCGAATTG 840
Qy 281 LeuGlyGlnGlyGlyPheGlyTyrValHisLysGlyValLeuProSerGlyLysGluVal 300
Db 841 TTAGGACACGCGGGTTCGGTTACGTGCACAAGAGTGTGTCTCTAGTGGAAAGAGTT 900
Qy 301 AlaValLysGlnLeuLysValGlySerGlyGlnGlyGluArgGluPheGlnAlaGluVal 320
```

Db 901 GCTGTGAAGCAGTTCAAAGTTGGGAGTGGTCAGGGAGAGAGGAGTTTCAGGCAGAGGTT 960
Qy 321 GluIleIleSerArgValHHisArgHHisLeuValSerLeuValGlyTyrCysIleAla 340
Db 961 GAGATCATCAGCAGAGTTCCACCACAGGCATCTGGTGTCTCTTGTGGTTATTGTCATGCC 1020
Qy 341 GlyAlaLysArgLeuLeuValTyrGluPheValProAsnAsnAsnLeuGluLeuHisLeu 360
Db 1021 GGTCGCAAAAGATTGCTTGTCTATGAGTTTGTCTTAACAACAATCTCGAGCTTCACCTC 1080
Qy 361 HisGlyGluGlyArgProThrMetGluTyrPheSerThrArgLeuLysIleAlaLeuGlySer 380
Db 1081 CATGCGAGGAGCGCCTACATGAATGGAGCACCAGATTGAAGATTGCTCTTGGATCT 1140
Qy 381 AlalysGlyLeuSerTyrLeuHisGluAspCysAsnProLysIleIleHisArgAspIle 400
Db 1141 GCTAAGGACTTCTTATCTTCATGAAGATTGCAATCTAAATCATTCACCGTGATATC 1200
Qy 401 LysAlaSerAsnIleLeuIleAspPheLysPheGluAlaLysValAlaAspPheGlyLeu 420
Db 1201 AAGGCTTCAACATATATTAGATTTCAGTTTGAAGCTAAGGTTGCTGATTTTGGTCTT 1260
Qy 421 AlalysIleAlaSerAspThrAsnThrHisValSerThrArgValMetGlyThrPheGly 440
Db 1261 GCTAAGATTGCTTCTGATACAAACACGATGATCAACACGCTGTGATCGGAACCTTTGGG 1320
Qy 441 TyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPheSer 460
Db 1321 TACTTGGCTCGGAATAGCTGCAAGCGGAAGCTCACGGAGAAGTCTGACGTTTCTTCA 1380
Qy 461 PheGlyValValLeuLeuGluLeuIleThrGlyArgArgProValAspAlaAsnVal 480
Db 1381 TTTGGCGTGTGCTTTTGGAGCTCAATTACTGGACGTCGACCCGTTGATGCCCAACAATGTC 1440
Qy 481 TyrValAspAspSerLeuValAspTyrAlaArgProLeuLeuAsnArgAlaSerGluGln 500
Db 1441 TATGTAGATGACAGTTAGTTGACTGGCAGCAGCAATGCTTACCGAGCATCTGAGCAA 1500
Qy 501 GlyAspPheGluGlyLeuAlaAspAlaLysMetAsnGlyTyrAspArgGluMet 520
Db 1501 GGAGACTTTGAGGGTTTACGTATGCAAAAGATGAATAATGGGTATGACAGAGAGAGATG 1560
Qy 521 AlaArgMetValAlaCysAlaAlaCysValArgHisSerAlaArgArgProArg 540
Db 1561 GCTCGATGGTGTGCTTGTGCGGCTTGTGTGCGCCATTGAGCTCGCCGAGACCTCGC 1620
Qy 541 MetSerGlnIleValArgAlaLeuGluGlyAsnValSerLeuSerAspLeuAsnGluGly 560
Db 1621 ATGAGCCAGATTGTCGTCGTTAGAGGAAATGTATCACTGCAGATCTTAAACGAAGGG 1680
Qy 561 MetArgProGlyGlnSerAsnValTyrSerSerTyrGlyGlySerThrAspTyrAspSer 580
Db 1681 ATGAGACCAGGTCAAAGCAATGTATACAGCTCATACGGAGGAAGCACCGATTATGACTCG 1740
Qy 581 SerGlnTyrAsnGluAspMetLysLysPheArgLysMetAlaLeuGlyThrGlnGluTyr 600
Db 1741 AGCCAGTACAATGAAGACATGAAGAAATTTAGGAAAAATGGCACTTGGAACTCAAGAGTAC 1800
Qy 601 AsnAlaThrGlyGluTyrSerAsnProThrSerAspTyrGlyLeuTyrProSerGlySer 620
Db 1801 AACGCCACGGGTGACTACAGTAATCCGACCATGACTATGACTGTACCCGCTCTGTTCA 1860
Qy 621 SerSerGluGlyGlnThrThrArgGluMetGluMetGlyLysIleLysArgThrGlyGln 640
Db 1861 AGCAGCAGGGGCAAAACCAACACCGCAATGGAGATGGGGAAGATTAAAGAAACCGGTGAG 1920
Qy 641 GlyTyrSerGlyProSerLeu 647
Db 1921 GGTATATGAGCCCTTCTCTT 1941

RESULT 3
ACF36549
ID ACF36549 standard; cDNA; 2189 BP.

XX ACF36549;
AC 18-DEC-2003 (first entry)
XX B. napus PERK1 receptor kinase encoding cDNA.
XX PERK1; transgenic; plant; proline-rich extensin-like receptor kinase;
KW wound; pathogen resistance; plant growth; seed production; gene; ss.
XX Brassica napus.
XX Key Location/Qualifiers
CDS 1..2189
FT /*tag= c
FT /note= "ABR82938; this protein contains the amino acids
FT corresponding to 5' and 3'UTR regions, though only the
FT relevant aa residues (ABR82937) is used in the invention"
FT 5'UTR 1..96
FT /*tag= a
FT CDS 97..2040
FT /*tag= b
FT /transl_except= (pos: 1332..1334, aa: Phe)
FT /product= "PERK1"
FT /note= "ABR82937; proline-rich Extensin-like Receptor
FT Kinase"
FT 3'UTR 2041..2189
FT /*tag= d
XX WO2003072763-A1.
XX 04-SEP-2003.
XX 28-FEB-2003; 2003WO-CA000274.
XX 28-FEB-2002; 2002CA-02373903.
XX 28-FEB-2002; 2002US-00086464.
XX (GORI/) GORING D.
XX (SILV/) SILVA N.
XX (HAFF/) HAFFANI Y Z.
XX Goring D, Silva N, Haffani YZ;
XX WPI; 2003-712727/67.
XX P-PSDB; ABR82937, ABR82938.
XX Producing a transgenic plant having an increased plant resistance, plant
XX growth or seed production comprises transforming a plant with a nucleic
XX acid molecule having a Proline-rich Extensin-like Receptor Kinase
XX activity.
XX Disclosure; Fig 1D; 123pp; English.
XX The invention relates to producing a transgenic plant having increased
XX plant height, number of branches, number of seed pods and/or seed
XX production compared to a non-transgenic plant, and/or quicker flowering
XX or later senescence compared to a non-transgenic plant. The method
XX involves transforming a plant with a vector including a Proline-rich
XX Extensin-like Receptor Kinase (PERK) nucleic acid molecule or a nucleic
XX acid molecule having PERK activity. The method, as well as the PERK
XX nucleic acid molecule and polypeptide, are useful in increasing plant
XX resistance to wounding and pathogens and in increasing plant growth and
XX seed production. The nucleic acid molecule and polypeptide may also be
XX used in producing transgenic plants or transgenic host cells. The present
XX sequence represents a cDNA encoding a B. napus PERK1 receptor kinase
XX polypeptide
XX Sequence 2189 BP; 544 A; 592 C; 494 G; 559 T; 0 U; 0 Other;
SQ

Alignment Scores:
Pred. No.: 5e-93 Length: 2189
Score: 3453.00 Matches: 647

Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-10-086-464-2 (1-647) x ACF36549 (1-2189)

1 MetSerSerAlaProSerProGlyThrGlySerProProSerProProSerAsnSerThr 20
97 ATGTCCTCGGCGCGCTCTCGGGGACTGGTTGCGCTCCATCTCCACCACATCAACTCCACA 156

21 ThrThrThrProProAlaSerAlaProProProThrThrProProSerSerProProPro 40
157 ACCACCACTCCTCTCCAGCTTCGCTCTCTCCACACACACCTTCTCTCTCCGCGC 216

41 ProSerThrThrProProSerProProProSerSerArgSerThrProSerAlaProPro 60
217 CCATCCACTATCCGACATCTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 276

61 ProSerProProThrProSerThrProGlySerProProProLeuProGlnProSerPro 80
277 CCATCTCCACCACTCCATCTACGCGGGATCTCCACCTCTCTCTCTCTCTCTCTCTCTCA 336

81 ProAlaProThrThrProGlySerProProAlaProValThrProProThrArgAsnPro 100
337 CCGCTCCCACTACGCGCGGATCTCCACCCGACCTGTCTCTCTCTCTCTCTCTCTCTCT 396

101 ProProSerValProGlyProProSerAsnProSerArgGlyGlySerProArgPro 120
397 CCACCTTCAGTCCCGAGGACCCGCTCAATCTCTCACGGAAGGAGGATCTCTCTCGACCT 456

121 ProSerSerProSerProProSerProSerAspGlyLeuSerThrGlyValVal 140
457 CCATCTCTCTCTCGCGCGCTCTCTCTCTCTCCGAGGCTTTATCAACAGGAGTGGTGGTG 516

141 GlyLeuAlaIleGlyValAlaLeuLeuValIleValThrLeuIleCysLeuLeuCys 160
517 GGAATCGCATCGGAGGAGTGGCTGTCTGTGTAGTAGTACTGTGATTTGTCTCTCTCTGT 576

161 LysLysLysArgArgArgAspGluGluAspAlaTyrTyrValProProProProPro 180
577 AAGAAGAACACCGAGAGACGAAAGATGCTTACTATGTCTCTCGCCACCTCTCTCTCT 636

181 GlyProLysAlaGlyGlyProTyrGlyGlyGlnGlnGlnTyrArgGlnGlnAsnAla 200
637 GGTCCCAAGCGGAGGACCTTACGGTGCACGACGACGACCAATGGCGGCAACAACGCA 696

201 ThrProProSerAspHisValThrSerLeuProProProProProProLysAlaProSerPro 220
697 ACACCCCGTCAGATCATGTCTGAGCTCACTACCCACCACCACTAAGGCTCCATCTCCA 756

221 ProArgGlnProProProProProProProProProProProProProProProPro 240
757 CCACGGCAACCTCTCCACCTCCACCCGCTTTTCATGACGACGCGCGGCTCCGAC 816

241 TyrSerAspArgProValLeuProProProProProGlyLeuValLeuGlyPheSerLys 260
817 TACTCGGACCGTCCAGTCTCTCTCTCCACCGTCTCCAGGGCTTGTGTAGGCTTCTCCAA 876

261 SerThrPheThrTyrGluGluLeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeu 280
877 AGCATTTTCATACGAGGAGCTAGTAGACCCCAATGGTTCTCCGAGGCGCAACTTG 936

281 LeuGlyGlnGlyPheGlyTyrValHisLysGlyValLeuProSerGlyLysGluVal 300
937 TTAGGACAAGCGGGTTCGGTTCGTGCACAAAGGTGTGTTCCTAGTGGGAAAGAGTT 996

301 AlaValLysGlnLeuLysValGlySerGlyGlnGlyGluArgGluPheGlnAlaGluVal 320
997 GCTGTGAACGAGTTGAAGTTGGAGTGGTTCAGGAGAGAGGAGTTTCAGGCAAGAGTT 1056

321 GluIleSerArgValHisArgHisLeuValSerLeuValGlyTyrCysIleAla 340

Db 1057 GAGATCATCAGCAGAGTTCCACACGAGCATCTGGTGTCTCTTGTGGTTATTGATCGCC 1116

Qy 341 GlyAlaLysArgLeuLeuValTyrGluPheValProAsnAsnLeuGluLeuHisLeu 360

Db 1117 GGTCCAAAAGATTGCTTGTCTATGAGTTTGTCTTAACAACAATCTCGAGCTTCACTC 1176

Qy 361 HisGlyGluGlyArgProThrMetGluThrSerThrArgLeuLysIleAlaLeuGlySer 380

Db 1177 CATCGCAGGAGCGCCTACAAATGAATGGAGCACCAGATTGAAGATTGCTCTTGGATCT 1236

Qy 381 AlalysGlyLeuSerTyrLeuHisGluAspCysAsnProLysIleIleHisArgAspIle 400

Db 1237 GCTAAAGAGCTTCTTATCTTCAAGAGATTGCAATCTTAATAATCAATTCACCGTGATC 1296

Qy 401 LysAlaSerAsnIleLeuIleAspPheLysPheGluAlaLysValAlaAspPheGlyLeu 420

Db 1297 AAGCTTCAACAATATTGATAGATTCAAGTTTGAAGCTAAGGTGTCTGATTTTGGTCTT 1356

Qy 421 AlalysIleAlaSerAspThrAsnThrHisValSerThrArgValMetGlyThrPheGly 440

Db 1357 GCTAAGATTGCTCTGATACAAACACGATGTATCAACACGCTGTGATGGGAACCTTTGG 1416

Qy 441 TyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPheSer 460

Db 1417 TACTTGGCTCGGAAATACGCTCAAGCGGAAGCTCACGGAGAAGTCTGACGTTTCTCA 1476

Qy 461 PheGlyValValLeuLeuGluLeuIleThrGlyArgArgProValAspAlaAsnVal 480

Db 1477 TTTGGCGTGTGCTTTTGGAGCTCATTTAGTGGAGTGCACCCGTTGATGCGCAACAATGC 1536

Qy 481 TyrValAspAspSerLeuValAspTrpAlaArgProLeuLeuAsnArgAlaSerGluGln 500

Db 1537 TATGTAGATGACAGCTTAGTTGACTGGCGACGACCATTTGCTTAACCGAGCATCTGAGCA 1596

Qy 501 GlyAspPheGluGlyLeuAlaAspAlaLysMetAsnAsnGlyTyrAspArgGluGluMet 520

Db 1597 GGAGACTTTGAGGTTTGTAGCTGATCAAGATGAATAATGGGTATGACAGAGAGAGATG 1656

Qy 521 AlaArgMetValAlaCysAlaAlaCysValArgHisSerAlaArgArgProArg 540

Db 1657 GCTCGCATGGTGTGCTGTGCTGCGGCTGTGTTCGCCATTCAGCTCGCCGAGACCTCGC 1716

Qy 541 MetSerGlnIleValArgAlaLeuGluGlyAsnValSerLeuSerAspLeuAsnGluGly 560

Db 1717 ATGAGCCAGATTGTCGCTGCGTTAGAGAGAAATGTATCATCTGTACAGATCTTAACGAGG 1776

Qy 561 MetArgProGlyGlnSerAsnValTyrSerSerTyrGlyGlySerThrAspTyrAspSer 580

Db 1777 ATGAGACCGGTCAAGCAATGTATACAGCTCATACGAGGAGGAGCAGCATATGACTCG 1836

Qy 581 SerGlnTyrAsnGluAspMetLysPheArgLysMetAlaLeuGlyThrGlnGluTyr 600

Db 1837 AGCCAGTACAAATGAAGACATGAAGAAAGTTTAGGAAAATGGCACTTGAACCTCAAGAGTAC 1896

Qy 601 AsnAlaThrGlyGluTyrSerAsnProThrSerAspTyrGlyLeuTyrProSerGlySer 620

Db 1897 AACGCCACGGGTGAGTACAGTAAATCCGACCACTGACTATGGAGTGTACCCGCTGCTTCA 1956

Qy 621 SerSerGluGlyGlnThrThrArgGluMetGlyLysIleLysArgThrGlyGln 640

Db 1957 AGCAGCGGCGCCAAACCAACACGCAATGGAGATGGGGAAGATTAAAGAGAACCGGTCA 2016

Qy 641 GlyTyrSerGlyProSerLeu 647

Db 2017 GGTTATAGTGGACCTTCTCTT 2037

RESULT 4

ACF36557

ID ACF36557 standard; DNA; 1959 BP.

XX ACF36557;

AC ACF36557;

XX ACF36557;

DT 18-DEC-2003 (first entry)

[illegible]

```
Db 1297 AACACTCATGTATCTACACGGCTGATGGAACTTTGGGTATTTGGCTCCGGAATATGCT 1356
Qy 448 AlaserGlyLysLeuThrGluLysSerAspValPheSerPheGlyValValLeuLeuGlu 467
Db 1357 GCAAGTGAAGAGCTCACAGAAAGTCTGACGTTTTCTCATTTGGCGTTGTACTTTTGGAA 1416
Qy 468 LeuIleThrGlyArgProValAspAlaAsnAsnValTyrValAspAspSerLeuVal 487
Db 1417 CTTATTACTGGAGGCGCCTGTTGATGGCAACATGCTATGTAGATGACAGCTTAGTT 1476
Qy 488 AspTrpAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGluGlyLeuAla 507
Db 1477 GACTGGGCACGACCATTCCTTAACCGAGCATCTCGAGGAAGGAGATTTTGAAGGTTTGCT 1536
Qy 508 AspAlaLysMetAsnAsnGlyTyrAspArgGluGluMetAlaArgMetValAlaCysAla 527
Db 1537 GATTCAAGATGGGTAATGAGTATGACAGAGAGAGATGGCTCGCATGTTGCTTGGCT 1596
Qy 528 AlaAlaCysValArgHisSerAlaArgArgProArgMetSerGlnIleValArgAla 547
Db 1597 GCGGCTTGTTGGCCATTGAGCTCGCCGACAGACCTCGCATGAGCCAGATAGTACGGCG 1656
Qy 548 LeuGluGlyAsnValSerLeuSerAspLeuAsnGluGlyMetArgProGlyGlnSerAsn 567
Db 1657 TTAGAAGGAATGATATCGCTGCTGATCTTAACGAAGGATGAGACCGGGTCAACGCAAC 1716
Qy 568 ValTyrSerSerTyrGlyLysThrAspTyrAspSerSerGlnTyrAsnGluAspMet 587
Db 1717 GTATACAGCTCATATGGAGGAAGCACACAGATATGACACGAGCCCAATACACGACGACATG 1776
Qy 588 LysLysPheArgLysMetAlaLeuGlyThrGlnGluTyrAsnAlaThrGlyGluTyrSer 607
Db 1777 ATAAAGTTAGAAAATGGCTCTTGGAACCTCAAGNATACGCACACACCGGCGAGTACAGT 1836
Qy 608 AsnProThrSerAspTyrGlyLeuTyrProSerGlySerSerSerGluGlyGlnThrThr 627
Db 1837 AATCCAACCACTGACTACGAGTGTACCCGCTCTGTTCAAGCAGTGAAGGTCAAGCCACA 1896
Qy 628 ArgGluMetGluMetGlyLysIleLysArgThrGlnGlyTyrSerGlyProSerLeu 647
Db 1897 CGAGAAATGGAGATGGGAAGATTAAAGAAACCCGCTCAAGGTTATAGTGGACCTCTCTT 1956

RESULT 5
ID AAC36968 standard; DNA; 2324 BP.
XX
AC AAC36968;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 15710.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
FN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 29-MAR-1999; 99US-0126284P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 05-MAY-1999; 99US-0132484P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140951P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144431P.
PR 19-JUL-1999; 99US-0144432P.
PR 19-JUL-1999; 99US-0144433P.
```



```
QY 368 etGluTrpSerThrArgLeuIysIleAlaLeuGlySerAlaIysGlyLeuSerTyrLeuH 388
Db 1302 TGGAGTTCTCCACTAGTTGGGTATCGCTTAGGTGCTCGAAAGGACTCGTTACCTTC 1361
QY 388 leGluAspCys----- 391
Db 1362 ACGAAGACTGTAAAGTTTTAAACATTCCACCATCTCATTTTCTTAACCAAGTTGCATAAAAC 1421
QY 392 -----Asn 392
Db 1422 AGAGAAAGCTCTGCTCTGACTAGTGTATCTTTTGGCTGAGAAATGGTGCAGGCCAT 1481
QY 393 ProLysIleIleHisArgAspIleLysAlaSerAsnIleLeuIleAspPheLysPheGlu 412
Db 1482 CCTCGGATCATTTACCGGACATCAAGCTGCAATATATCTTTGGACTTCAACTTTGAT 1541
QY 413 Ala----- 413
Db 1542 GCTATGGTGAATAACTAGTAGTTCGCAATTCATCTACGGTTTTTTGTTTAAGACTACATTGA 1601
QY 414 -----LysValAlaAspPheGlyLeuAlaLysIleAlaSerAs 426
Db 1602 TGACATTTTGCATTTGTTTATTTCAGGTGGCTGATTTTGGATTAGCTAAGTTAAACATCTGA 1661
QY 426 pThrAsnThrHisValSerThrArgValMetGlyThrPhe-Gly----- 440
Db 1662 TAACAACACTCATGTATCTACTCGGTGTGATGGGAACCTTTTCGGGTAAAGCGTTTACCCTAT 1721
QY 440 ----- 440
Db 1722 GATAAGATTGTCGTGACACTCAAGAAACATAACCTTTGTAGACTAATCTACTTTGCTTC 1781
QY 441 -----TyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrG 454
Db 1782 TTTCCACAAACATGTGTAGATATCTAGCTCCAGAAATATGCTCAAGCGGTAAATTAACCG 1841
QY 454 luLysSerAspValPheSerPheGlyValValLeuLeuGluIleThrGlyArgArgP 474
Db 1842 AGAAATCCGATGTTTTCTCTTACGGAGTATGTATTATCGAACTTATACTGGAAACGAC 1901
QY 474 roValAspAlaAsnValTyrValAspArgGluMetAlaArgMetValAlaCysAlaAla 489
Db 1902 CGGTTGAT---AATAGCATCACCATTGACGACACCTTAGTAGATTGGGTATTCATGCAATG 1958
QY 489 ----- 489
Db 1959 TAACATATGTATCGGTATATATGTTTTTTCGGCTTTTTCGGCTACTAATGATCATGAATA 2018
QY 490 ---AlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGluGlyLeuAlaAsp 508
Db 2019 CAGGCTCGGCTCTTATCGCTCGCGCTAGAGATGGAACCTTTAATGAGCTCGCAGAT 2078
QY 509 AlaLysMetAsnAsnGlyTyrAspArgGluGluMetAlaArgMetValAlaCysAlaAla 528
Db 2079 GCGAGCTTGAAGGCACTACAAACCCGCAAGAAATGGCTCGAATGGTGAATGTTGCGCGCT 2138
QY 529 AlaCysValArgHisSerAlaArgArgProArgMetSer----- 542
Db 2139 GCTAGCATTCGTCAATTCGGGGGTAAACGTCCTCAAGATGAG-CCAGGTGAATCAAAATTA 2197
QY 543 -----GlnIleVal 545
Db 2198 TAACATAAAGTCTATTTTTTGTGAGAGAATAACAAACAATGTGTGGTTTTTCAGATAGTA 2257
QY 546 ArgAlaLeuGluGlyAsnValSerLeuSerAspLeuAsnGluGlyMetArgProGlyGln 565
Db 2258 AGAGGCTTAGAAGGAGAGATGCTCTTAGATGCTTTAAACGAGGTGTGAAGCCAGGACAC 2317
QY 566 SerAsnValTyrSerTyrGlyGlySerThrAspTyrAspSerSerGlnTyrAsnGlu 585
Db 2318 AGTACGTTTACGGGTCAATTCGGGCAAGCTCGGATTATAGTCAGACATCTTACAATGCA 2377
```

```
QY 586 AspMetLysLysPheArgLysMetAlaIleuGlyThrGlnGluTyr-----AsnAla 602
Db 2378 GACATGAAGAAATTCAGACAGATAGCTTTGTCGAGCCAAAGAAATCCCAAGTCAGTACTGT 2437
QY 603 ThrGlyGluTyrSerAsnProThrSerAspTyrGly 614
Db 2438 GAAGGAACATCTAGTAATGATTCTTAGAGATATGGGA 2473

RESULT 10
ACF36551
ID ACF36551 standard; DNA; 2520 BP.
XX
AC ACF36551;
XX
DT 18-DEC-2003 (first entry)
XX
DE Arabidopsis PERK1 receptor related protein encoding DNA.
XX
KW PERK1; transgenic; plant; proline-rich extensin-like receptor kinase;
KW wound; pathogen resistance; plant growth; seed production; gene; ds.
XX
OS Arabidopsis thaliana.
XX
PN WO2003072763-A1.
XX
PD 04-SEP-2003.
XX
PF 28-FEB-2003; 2003WO-CA000274.
XX
PR 28-FEB-2002; 2002CA-02373903.
XX
PR 28-FEB-2002; 2002US-00086464.
XX
PA (GORI/) GORING D.
PA (SILV/) SILVA N.
PA (HAFF/) HAFFANI Y Z.
XX
PI Goring D, Silva N, Haffani YZ;
XX
DR WPI; 2003-712727/67.
DR P-PSDB; ABR82940.
XX
PT Producing a transgenic plant having an increased plant resistance, plant
PT growth or seed production comprises transforming a plant with a nucleic
PT acid molecule having a Proline-rich Extensin-like Receptor Kinase
PT activity.
XX
PS Disclosure; Fig 12; 123pp; English.
XX
CC The invention relates to producing a transgenic plant having increased
CC plant height, number of branches, number of seed pods and/or seed
CC production compared to a non-transgenic plant, and/or quicker flowering
CC or later senescence compared to a non-transgenic plant. The method
CC involves transforming a plant with a vector including a Proline-rich
CC Extensin-like Receptor Kinase (PERK) nucleic acid molecule or a nucleic
CC acid molecule having PERK activity. The method, as well as the PERK
CC nucleic acid molecule and polypeptide, are useful in increasing plant
CC resistance to wounding and pathogens and in increasing plant growth and
CC seed production. The nucleic acid molecule and polypeptide may also be
CC used in producing transgenic plants or transgenic host cells. The present
CC sequence represents a PERK1 polypeptide related protein encoding genomic
CC DNA from A. thaliana (Accession NO. AAD15491)
XX
SQ Sequence 2520 BP; 717 A; 552 C; 538 G; 713 T; 0 U; 0 Other;
```

Alignment Scores:

Pred. No.:	3.58e-36	Length:	2520
Score:	1504.00	Matches:	367
Percent Similarity:	53.47%	Conservative:	80
Best Local Similarity:	43.90%	Mismatches:	146
Query Match:	43.56%	Indels:	224
DB:	9	Gaps:	22

US-10-086-464-2 (1-647) x ACF36551 (1-2520)

Db 987 CCTGTGCTGTGCTCT-----CACAGTGATGCTTCCAACTTAACCGGT 1028
Qy 244 ArgProValLeuProProSerProGlyLeuValLeuGlyPheSerLysSerThrPhe 263
Db 1029 CGAACTGTATACCGTCACTCAA---CCTGCAACTCTTGGTCACAAAGACACTTTC 1085
Qy 264 ThrTyrGluGluLeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuGlyGln 283
Db 1086 ACATACGATGAATGCTCCATTGCAACAGAGGTTTCGCTCAGTCAAATTTGCTAGGACAA 1145
Qy 284 GlyGlyPheGlyTyrValHisGlyValLeuProSerGlyLysGluValAlaValLys 303
Db 1146 GGAGGATTTGGTATGTTTCATAAAGGAGTCTGCTAGTGGCAAGAAAGTTGCAAGTGAAG 1205
Qy 304 GlnLeuLysValGlySerGlyGlnGlyGluArgGluPheGlnAlaGluValGluLeuLe 323
Db 1206 AGTCTTAAACTTGGNAGTGGACAGGGGAAACGAGTTTCAAGCAGAGTTGATATCAT 1265
Qy 324 SerArgValHisArgHisLeuValSerLeuValGlyTyrCysIleAlaGlyAlaLys 343
Db 1266 AGCGGTGTCCATCATCGTCACTCGTCTCTGTTGGATATTGCATCTCTGGTGGTCAA 1325
Qy 344 ArgLeuLeuValTyrGluPheValProAsnAsnLeuGluLeuHisHis----- 361
Db 1326 AGACTTTTGGTTATGAGTTTATACCTTAACAACACTCTTGAATTTCAATTCATGGTACA 1385
Qy 361 ----- 361
Db 1386 TTCATCTAACAGATGTTTCTGTATTAAACAAACCTTTAAGTATGTTCTCTTTAAT 1445
Qy 362 -----GlyGluGlyArgProThrMetGluTrpSerThrArgLe 374
Db 1446 CAGGAACATGATTGAAATTTTCAGGAAGAGGTCGTCGGTTTTGGATTGGCTACAGAGT 1505
Qy 374 uLysIleAlaLeuGlySerAlaLysGlyLysSerTyrLeuHisGluLysPhe----- 391
Db 1506 GAAGATTGTCATTGGGATCAGCTAGAGGCTTGCATATTGTCATGAAGACTGTAAAGAAAT 1565
Qy 391 ----- 391
Db 1566 CTTTATCTCACATATTGTCATCAGTTTCTATCTCGTCTCTACAAATATTGAAAGATTGT 1625
Qy 392 -----AsnProLysIleIleHisArgAspIleLysAlaSerAs 404
Db 1626 ATATGTTATACATATTATAGTCAACCTCGCATTTATCCACAGATATCAAGCTGCAAA 1685
Qy 404 nIleLeuIleAspPheLysPheGluAla----- 413
Db 1686 CATTCCTTCTGATTTTCAGTTTGTAGAC-CAAGGTATGTGTATATATATCGACTCTTGTAC 1744
Qy 413 ----- 413
Db 1745 TACTTTTACTTTCATTTGCTCTCATTTTGTGTTTCCAAATCTGTGTCGATGTTGATCAGT 1804
Qy 414 -----LysValAlaAspPheGlyLeuAlaLysIleAlaSerAspTh 427
Db 1805 CTTATTGTGTAATATATGTCAGGTGGCAGATTTTGGATTGGCTAAGCTATCTCAAGACAA 1864
Qy 427 rAsnThrHisValSerThrArgValMetGlyThrPhe-Gly----- 440
Db 1865 CTATACTCATGTCTCCACTCGCTCATGGGAACCTTTTGGGTAAAGCAGCTTTGTAATAATGT 1924
Qy 440 ----- 440
Db 1925 CTCAACTCATCCACATTAATTAGTTTCTTTCATCTGTTTAAACATTTTCTTGGATTCA 1984
Qy 441 --TyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPheS 460
Db 1985 GATACTTAGCTCCAGAGTATCATCAAGCGGAAGTATCCGACAAATCTGATGTTTCT 2044
Qy 460 erPheGlyValValLeuLeuGluLeuIleThrGlyArgArgProValAspAlaAsnV 480
Db 2045 CATTTGGAGTAATGCTTCTTGAGCTCATAAACCGGAAGACCTCTCTCGATCTAACTGGA- 2103

Qy 480 alTyrValAspAspSerLeuValAspTyr----- 489
Db 2104 --GAAATGGAAGATAGCTTGGTAGATTGGGTAAAGTCGGTCCCCGCTCTTCGGTTTACTT 2161
Qy 489 ----- 489
Db 2162 GTTTTAATCCCCAAACACTTTTCCAAAGCAAAAAACAGAAACAAATCTTACTATTGTTGTTC 2221
Qy 490 --AlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGluGlyLeuAlaAspA 509
Db 2222 AGGCAAGCGCTTGTGTTGAAAGCAGCTCAAGATGGAGATTACAAACCAATTGGCTGATC 2281
Qy 509 laLysMetAsnAsnGlyTyrAspArgGluGluMetAlaArgMetValAlaCysAlaAlaA 529
Db 2282 CACGCTAGAGCTAAACTACAGTCAATCAAGAGATGGTTCAATGGCTTCTTGTGTCAGCTG 2341
Qy 529 laCysValArgHisSerAlaArgArgProArgMetSerGlnIle----- 544
Db 2342 CAGCAATCAGACATTCAGCAAGAGAGCGGCTTAAGATGAGCCAGGT-TCAAAAAACTCATA 2400
Qy 544 ----- 544
Db 2401 CCACTTGTGTTCTTATGTTATATTTTACTCAAAATTAATCTTGATGATAAATGTGAC 2460
Qy 545 -----ValArgAlaLeuGluG 550
Db 2461 ATACTAATGAATCTTGAACATGTGTATGTTAAATGAAAGATTGTACGACACTAGAG 2520
Qy 550 lyAsnValSerLeuSerAspLeuAsnGluGlyMetArgProGlyGlnSerAsnValTyrS 570
Db 2521 GAGATATGTCAATGAGTATCTAAGTGGGAAACAGACAGGACAAAGCAGCTACTTGA 2580
Qy 570 erSerTyrGlyGlySerThrAspTyrAspSerSerGlnTyrAsnGluAspMetLysLysP 590
Db 2581 GCCCCGGAGCGTGAGCTCAGAGTATGCGCAAGCTCGTACACGCGACAGATGAAAAAGT 2640
Qy 590 heArgLysMetAlaLeuGlyThrGlnGluTyrAsnAlaThrGlyGluTyrSerAsnPro 610
Db 2641 TCAGAAACTGGCGTTAGAGNATAAGNATATCAAGCAGT--GAATATGTTGGA---A 2694
Qy 610 hrSerAspTyrGlyLeuTyrProSerGlySerSerGluGlyGlnThrThrArgGluM 630
Db 2695 CAAGTGAGTATGGCTTAAACCTTCTGCTTCAAGTAGTGAA-----GAAA 2739
Qy 630 etGluMetGlyLysIleLysArgThrGlyGln 640
Db 2740 TGAATAGAGGCTCAATGAAACGCAATCCTCAG 2771

RESULT 12

ACF36552
ID ACF36552 standard; DNA; 2820 BP.

AC ACF36552;

XX 18-DEC-2003 (first entry)

DT Arabidopsis PERK1 receptor related protein encoding DNA.

DE PERK1; transgenic; plant; proline-rich extensin-like receptor kinase;
KW wound; pathogen resistance; plant growth; seed production; gene; ds.

XX Arabidopsis thaliana.

XX WO2003072763-A1.

PN 04-SEP-2003.

PD 28-FEB-2003; 2003WO-CA000274.

PF 28-FEB-2002; 2002CA-02373903.

PR 28-FEB-2002; 2002US-00086464.

XX


```
Qy 391 ----- 391
Db 1566 CTTTATCTCACATATTTCATCAGTTTCTATCTCGCTCTCTACAATAATTTTGAAGATTGT 1625
Qy 392 -----
Db 1626 ATATGTTTACATCAATATATAGTGCACCTTCGATTTATCCACAGATATCAAAAGCTGCAAA 1685
Qy 404 nileLeuileAspPheLysPheGluAa- 413
Db 1686 CATCTCTTGATTTTCAGTTTTCAGAC-CAAGGTATGTGTGTATATATCGACTCTTGTATC 1744
Qy 413 ----- 413
Db 1745 TACTTTTACTTTTCATTTGCTCTCATTTTGTTCATCTGTGTCGATGTGTATCAGT 1804
Qy 414 -----
Db 1805 CTTATGTGTAAATATATGTCAGGTGGCAGATTTTGGATTTGGTAAAGCTATCTCAAGACAA 1864
Qy 427 xAsnThrHisValSerThrArgValMetGlyThrPhe-Gly- 440
Db 1865 CTATCTCATGTCTCCACTCGCTCGGTCAATGGGAACCTTTGGGTAGCAGCTTTGTAATATGT 1924
Qy 440 ----- 440
Db 1925 CTCAACTCATCCACACTTATTAGTTTCTTTTCACTTTGTTTAAACATTTTCTTGGATTCA 1984
Qy 441 --TyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPheS 460
Db 1985 GATACTTAGCTCCAGAGTATCATCAAGCGGAAAGTTATCCGACAAAATCTGATGTTTCT 2044
Qy 460 erPheGlyValLeuLeuGluLeuileThrGlyArgProValAspAlaAsnAsnV 480
Db 2045 CATTGGAGTATGCTTCTGAGCTCATACCGGAAGACCTCTCTGGATCTAACTGGA- 2103
Qy 480 alTyrValAspSerLeuValAspTrp- 489
Db 2104 --GAAATGGAAGATAGCTTGTAGATTGGTAAAGTTCGGTCCCGCTCTTCGGTTTACTT 2161
Qy 489 ----- 489
Db 2162 GTTTAATCCCAAAACACTTTCCAAAGCAAAACAGAAACAATCTTACTATTTGTTGTC 2221
Qy 490 --AlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGlyLeuAlaAspA 509
Db 2222 AGCAAGCCCTTTGTTGTTGAAGCAGCTCAAGATGGAGATTACAACCAATTTGCTGATC 2281
Qy 509 laLysMetAsnAsnGlyTyrAspArgGluGluMetAlaArgMetValAlaCysAlaAlaA 529
Db 2282 CACGCTAGAGCTAAACTACAGTCAATCAAGAGATGGTTCAATGGCTTCTTGTGCAGCTG 2341
Qy 529 laCysValArgHisSerAlaArgArgProArgMetSerGlnile- 544
Db 2342 CAGCAATCAGACATTCAGCAAGAAGACGGCCTAAGATGAGCAGGT-TCAAAAACTCATA 2400
Qy 544 ----- 544
Db 2401 CCACCTGTTGTTCTATTGTTATATTTTACTCAAAATTAATCTTGATGATAAATGTGAC 2460
Qy 545 -----
Db 2461 ATACTAATGAATCTTGAACATGTTATGTTGTAATGAAAGATTGTACGACACTAGAG 2520
Qy 550 lyAsnValSerLeuSerAspLeuAsnGluGlyMetArgProGlyGlnSerAsnValTyrS 570
Db 2521 GAGATATGTCATGATGATCTAAGTGGGGAACAAGACAGGACAAAGCAGCAGTACTTGA 2580
Qy 570 erSerTyrGlyLysSerThrAspTyrAspSerSerGlnTyrAsnGluAspMetLysLysP 590
Db 2581 GCCCGGAGCGTGAAGTACAGAGTATGACGCAAGCTCGTACACGGCAGACATGAAGAAGT 2640
Qy 590 heArgLysMetAlaLeuGlyThrGlnGluTyrAsnAlaThrGlyGluTyr-SerAsnProT 610
```

```
Db 2641 TCAAGAAACTGGCGTTAGAGAATAAAGAATATCAAGCAGT- 2694
Qy 610 hrSerAspTyrGlyLeuTyr-ProSerGlySerSerGluGlyGlnThrThrArgGluM 630
Db 2695 CAAGTGAATGGCTTTAAACCCCTTCTGCTTCAAGTAGTGA- 2739
Qy 630 etGluMetGlyLysIleLeuArgThrGlyGln 640
Db 2740 TGAATAGAGGCTCAATGAACGCAATCCTCAG 2771
RESULT 13
AAF77095
ID AAF77095 standard; DNA; 2880 BP..
AC AAF77095;
XX
DT 17-MAY-2001 (first entry)
XX
DE Arabidopsis gene #1.
XX
KW Proline-rich extensin-like receptor kinase; PERK; resistance; plant; ds.
XX
OS Arabidopsis thaliana.
XX
PN WO200114563-A1.
XX
PD 01-MAR-2001.
XX
PF 18-AUG-2000; 2000WO-CA000966.
XX
PR 19-AUG-1999; 99US-0149466P.
XX
PR 13-OCT-1999; 99US-0159122P.
XX
XX (GORI/) GORING D.
XX (SILV/) SILVA N.
XX
XX Goring D, Silva N;
XX
XX WPI; 2001-244305/25.
XX
PT New proline-rich, extensin-like receptor kinase nucleic acids and
PT polypeptides useful for increasing plant wounding or pathogen resistance,
PT or for producing transgenic plants with increased wounding or pathogen
PT resistance.
XX
PS Example; Fig 11; 91pp; English.
XX
CC The present invention relates to proline-rich extensin-like receptor
CC kinase (PERK). The PERK nucleic acids and polypeptides are useful for
CC increasing the resistance of plants to wounding and pathogens. These are
CC also useful for producing transgenic plants with increased wounding and
CC pathogen resistance compared with a wild type plant, as well as in assays
CC for identifying and developing compounds to inhibit and/or enhance
CC polypeptide function directly
XX
SQ Sequence 2880 BP; 814 A; 693 C; 569 G; 804 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 1.5e-27 Length: 2880
Score: 1209.50 Matches: 325
Percent Similarity: 43.85% Conservative: 92
Best Local Similarity: 34.17% Mismatches: 160
Query Match: 35.03% Indels: 376
DB: 4 Gaps: 26
US-10-086-464-2 (1-647) x AAF77095 (1-2880)
Qy 2 SerSerAlaProSerProGlyThrGlySerProSerProSerAsnSerThrThr 21
Db 98 AACTCAGCTCTTCCACCTGTTGATTCCTCTCCTAGTCCACGAGCTGATTCATCA 157
Qy 22 Thr-----ThrProProAlaSer----- 28
```


Qy 130 -----SerSerAspGlyLeuSerThrGly-----ValVal 139
Db 695 CCAAAACCAAGCAATGGAGATGGCGGTGGCGGTATCATAGGGAAGACTATG 754
Qy 140 ValGlyIleAlaIleGlyValAlaLeuLeuValIleValThrLeuIleCysLeuLeu 159
Db 755 GTTGGTATGGCTAGCCGGTTTGCATCATAGCGCTTATAGCGGTGTGTCTTAGTG 814
Qy 160 CysIysIysIysArgArgAspGluGluAspAlaIle-----TyrValPro 175
Db 815 AGAAGAAAGAAAAGAGA-----AACATTGATAGTATATCACTCACAGTACTGGCA 868
Qy 176 ProPro----- 177
Db 869 CATCCCAATTTCTCTGTTAAATCAGGTTTAAAAATCTCACCTTTATCTCTCTCAT 928
Qy 177 ----- 177
Db 929 CTTCTATGTGCTTGAATCATCTCTGACTATCTTTGCTTTTGATGTAGATGGATTCTTA 988
Qy 178 -----ProProGlyProIysAlaGlyGlyProIysGlyGlyGlnGlnGln 194
Db 989 TACGGTCAAGATCCAGGTAAGGATACCTCTCTGCTCAATATGTTCAATGTATATAACAAT 1048
Qy 195 TrpArgGlnGlnAsnAlaThrProProSerAspHisValValThrSerLeuProPro 214
Db 1049 TCACAGCAACAACTCTCTATGGAAACAGTTATGTGTACAGCTGGT----- 1096
Qy 215 ProIysAlaProSerProArgGlnProProProProProProProProPheMetSer 234
Db 1097 -----GGTGGTTATCTCATCATCA-----ATGCCAA 1123
Qy 235 SerSerGlyGlySerAspTyrSerAspArgProValLeuProProSerProGlyLeu 254
Db 1124 TCAAGTGCACACCTGAC-----TCTGCT 1147
Qy 255 ValLeuGlyPheSerLysSerThrPheThrTyrGluGluLeuAlaArgAlaThrAsnGly 274
Db 1148 ATACTCGGAAGTGCACAGCTCATTTACAGTACGAGAGCTTGTGAGATAAACAAGG 1207
Qy 275 PheSerGluAlaAsnLeuGlyGlnGlyPheGlyTyrValHisLysGlyValLeu 294
Db 1208 TTTGCTGCCAAAACATTTCTTGGAGAGCGGATTTGGATGTGTCTATAAAGGTACATG 1267
Qy 295 ProSerGlyLysGluValAlaValLysGlnLeuLysValGlySerGlyGlnGlyArg 314
Db 1268 CAGGATGGTAAAGTTGTTGCGTTAAAGCAGCTTAAAGCTGGAAGTGGACAAGGTGACCGT 1327
Qy 315 GluPheGlnAlaGluValGluIleSerArgValHisHisArgHisLeuValSerLeu 334
Db 1328 GAATTCAAAGCAGAGGTTGAGATCATCAGCCGCTTCATCATCGCCATTTGCTCTCTG 1387
Qy 335 ValGlyTyrCysIleAlaGlyAlaLysArgLeuLeuValTyrGluPheValProAsnAsn 354
Db 1388 GTTGGTTACTGCATTTACAGCAGCATAGATTGCTTATCTATGATGATGTTTCTTAATCAA 1447
Qy 355 AsnLeuGluLeuHisLeuHisGlyGlu----- 363
Db 1448 ACCTTGGAGCATCATTTGTCATGGTGAGTGACTTGTACCATTTCTCGTTATAGATAAGACT 1507
Qy 364 -----GlyArgProTh 367
Db 1508 TTTTTTTAGCTTTACGTGTTAGACTGCTCGCTTTACGCTTTTAGGAAAGGTTTGGCCAGT 1567
Qy 367 rMetGluTrpSerThrArgLeuValIleAlaLeuGlySerAlaLysGlyLeuSerTyrLe 387
Db 1568 TTTAGATGGTCTAAGAGAGTCCGATCGCTATAGGATCAGCCAAAGGTTGGCATATCT 1627
Qy 387 uHisGluAspCysAsn----- 392
Db 1628 TCACGAAGACTGTAA-GTAATGCCTTCACATTTCTTAGTTGTGTGCTTTGTTATGTCAC 1686

Qy 393 -----ProLy 394
Db 1687 TTCATAGTTTAAACAGAAAGCAAAATCATATCCTTGTTTTATTTTTCAGGTCTACCGAA 1746
Qy 394 siIleIleHisArgAspIleLysAlaSerAsnIleLeuIleAspPheLysPheGluAlaLy 414
Db 1747 AATCATTTCCAGAGATATAAGTCAGCAATATTTCTCTAGATGATGAATATGAAGCTCA 1806
Qy 414 s----- 414
Db 1807 -GGCAATAATGAATCCTCTCTTTTCGTTAAATCTATCTTATGACTGTAAAGTTTAGTTA 1865
Qy 415 -----ValalaAspPheGlyLeuAlaLysIleAl 424
Db 1866 ATGAGACTGTTCTGTTTTTTTGGATGTTAGTTGCTGATTTTGGACTTGTGACTCA 1925
Qy 424 aSerAspThrAsnThrHisValSerThrArgValMetGlyThrPhe-Gly 440
Db 1926 TGATACAACACAACTCATGTTTCAACTCGGGTTATGGGAACCTTCGGGTAAAGCAACAT 1985
Qy 440 ----- 440
Db 1986 TCATCACAAACTCTACTCTCCAAACTCGACCTTATTGATCCAATGCTCATGAAAAGTTTG 2045
Qy 440 ----- 440
Db 2046 TTATATATGCTTGAGGCAACAAATTTGGATCAAACTGGAATCTTTTATTTGATCGTATGGCT 2105
Qy 441 -----TyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuT 453
Db 2106 GCATGACATGTTTGTGTTAAGGTACCTAGCGCGGAATATGCATCAAGTGGNAATTTGA 2165
Qy 453 hrGluLysSerAspValPheSerPheGlyValValLeuLeuGluLeuIleThrGlyArgA 473
Db 2166 CTGATAGATCCGATGATTTCTCATTCGGGGTTGTTCTCTTAGAGTCTTAACATGACGGA 2225
Qy 473 xGProValAspAlaAsnAsnValTyrValAspAspSerLeuValAspTrp----- 489
Db 2226 AACCAGTTGACACAGCTCAGCCTCTAGGAGAAGAGAGTTGGTTGAATGGGTAAAGATCC 2285
Qy 489 ----- 489
Db 2286 AACTTTCAACATTTCTTAATAATAGTATGAGTTGGCCCTAGTATATATATAGTACTTA 2345
Qy 490 -----AlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGlu 504
Db 2346 TAAATGAACCTCACAGGCGCGCCGCTCTCTCAAGCCATTGAGACCGGAGATTTAAGC 2405
Qy 505 GlyLeuAlaAspAlaLysMetAsnAsnGlyTyrAspArgGluGluMetAlaArgMetVal 524
Db 2406 GAACGTGATTGATACACCGCTTGAAAAGCGTTATGTGGAGCATGAAGTCTTCAGATGATC 2465
Qy 525 AlaCysAlaAlaCysValArgHisSerAlaArgArgProArgMetSerGln--- 543
Db 2466 GAGACAGCGCTGCATGTGTTTAGACATTTCTGGTCCAAAACGCCACCGATGGTTCAAGTA 2525
Qy 543 ----- 543
Db 2526 ATTCGTACTAACCAAAAGTCCAAAGCTCCCATATATATAGTAACAAAGTATTCTCACACTCT 2585
Qy 544 -----IleValArgAlaLeuGluGlyAsnValSerLeu 554
Db 2586 GAAACCTTATCTACTCTTCGAAATAAGGTTGTGAGACATTGGACTCGGACGGAGACTCG 2645
Qy 555 SerAspLeuAsnGluGlyMetArgProGlyGlnSerAsnValTyrSerSerTyrGlyGly 574
Db 2646 GGAGATATTACCAACCGAATCAAAATTTGGSCAA----- 2678
Qy 575 SerThrAspTyrAspSerSerGlnTyrAsnGluAspMetLysLysPheArgLysMetAla 594
Db 2679 AGCACAACTTATGACTCAGGCGCAATACAATGAAGACATATTGAAATTCAGGAAATGGCG 2738
Qy 595 LeuGlyThrGlnGluTyrAsnAlaThrGlyGluTyrSerAsnProThrSerAspTyrGly 614


```
Db 1102 GGGCGTAAATGCTGTGACGAACACGACATGTTATACACA---ATGCCACCAACCAATCCA 1158
Qy 216 sAlaProSerProProArgGlnProProProProProProProProProProProPheMetSerSerSe 236
Db 1159 TGCTAAA-----TATATATCTAG 1176
Qy 236 rGlyGlySerAspTyrSerAspArgProValLeuProProSerProGlyLeuValLe 256
Db 1177 TGGAGTTGTTGATACGAAGAGAAC-----AATTCTGTTGCAAAAAACAT 1221
Qy 256 uGlyPheSerLysSerThrPheThrTyrGluGluLeuAlaArgAlaThrAsnGlyPheSe 276
Db 1222 TTCATGCACTCTGGAATGTTCTCTCAGAGAAGCTTTCAAAGCAACTGGTGATTTTC 1281
Qy 276 rGluAlaAsnLeuLeuGlyGlnGlyPheGlyTyrValHisLysGlyValLeuProSe 296
Db 1282 AGAGGAGAACCTTTTGGGAGAGGCGGTTTCGGATATGTTCAAAAGGAGTGTGAAAAA 1341
Qy 296 rGlyLysGluValAlaValLysGlnLeuLysValGlySerGlyGlnGlyGluArgGluPh 316
Db 1342 CGGCACAGAAGTTGCGGTGAAGCAGCTGAAGATTGGGAGCTATCAAGGGGAAAGAGAAAT 1401
Qy 316 eGluAlaGluValGluLleIleSerArgValHisHisArgHisLysValSerLeuValG1 336
Db 1402 CCAAGCTGAGGTTGACACAAATCAGTAGGGTTTCATCAAGCAGCCTCGTTTCATTGGTTGG 1461
Qy 336 YTyrcysIleAlaGlyAlaLysArgLeuLeuValTyrGluPheValProAsnAsnAsnLe 356
Db 1462 TTATTGCGTTAATGGAGATAAAGACTCTTGTTTACGAGTTTGTTCTTAAGATACCTT 1521
Qy 356 uGluLeuHisLeuHisGly----- 362
Db 1522 GGAGTTCACACTGCATGTGATAAATAGATATATGATTTCATCCTTTTTCATTGTTCTTGA 1581
Qy 363 -----GluGlyArg---ProThrMetG1 369
Db 1582 GTTCATTATGTTAGTATGTGAGAAATATGTTGTAGAGAACAGGAAAGCGTTGTGA 1641
Qy 369 uTPSerThrArgLeuLysIleAlaLeuGlySerAlaLysGlyLeuSerTyrLeuHisG1 389
Db 1642 ATGGAAATAGGCTCAGGATTCGTGTAGGAGCAGCAAAAGGATTAGCTTATCTTCATGA 1701
Qy 389 uAspCys----- 391
Db 1702 GGATTGTGAGTTGTTCTCTTCATAATGGGAATGACAATGGCCTTTTGGGTTTTCGAC 1761
Qy 392 -----AsnProLysIleIleHisArgAsp11 400
Db 1762 ACTGATATGATTTCTGTGTGTTGTAATGCGAGCGAGTCCAACTATAATTCACCGTGATAT 1821
Qy 400 eLysAlaSerAsnIleLeuIleAspPheLysPheGluAlaLys----- 414
Db 1822 CAAGCAGCTAATATCTCTTCATGATTCCAAATTTGAGGCAAAAGGTGATCTGCTCTTTAAT 1881
Qy 414 ----- 414
Db 1882 CTTATCAAAAGTTTGGTTTTAGAACAGAGATTTTGCCCATTTTCGGTTTATAGTCACACCAT 1941
Qy 415 -----ValAlaAs 417
Db 1942 TTTGTTCTATTAAACAAGATTACATGTAGCTTAGAGTTCTTGTTCTCTTCAGGTCCTGGA 2001
Qy 417 pPheGlyLeuAlaLysIleAlaSerAspThrAsn-----ThrHisValSerThrAr 434
Db 2002 CTTTGGACTAGCCAGTTTCTCAGACACCAATTCATCATCTCATATCTCTACTCTG 2061
Qy 434 gValMetGlyThrPheGly----- 440
Db 2062 AGTGGTAGGAACCTTTGCGGTAAAAACACCATCCATCCATGCTTTATATGTTGTGTCATTGT 2121
Qy 440 ----- 440
```

Search completed: April 24, 2004, 22:22:00
Job time : 666 secs

```
Db 2122 GTTTAAAAATATTAAATTTACGGTTTCAAGTTTCAAGTTTCAAGTTTCTTCCTT 2181
Qy 440 ----- 440
Db 2182 GTTGTATCGTAAATCCAGATCAAAAGATTTATTGATTACTAAATGCTTGTGCACCGTCT 2241
Qy 441 -----TyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuTh 453
Db 2242 ATTTGGTATGATTTAAACAGATACATGGCTCCAGAATACCGCTCCAGTGTGAAGTAAAC 2301
Qy 453 rGluLysSerAspValPheSerPheGlyValValLeuLeuLeuLleThrGlyArgAr 473
Db 2302 TGATAAATCAGATGATATCTTCTTGGGTGCTGCTTCTAGAACTCATCACTGGACGTC 2361
Qy 473 gProValAspAlaAsnAsnValTyrValAspAspSerLeuValAspTyr----- 489
Db 2362 ATCAATTTTCGCCAAAGATTTCTCCACAAACAGAGTTTGTAGACTGGGTAAAGTCAAAG 2421
Qy 489 ----- 489
Db 2422 TACATGATGATGATGATGATACCAATAGTGTCTTTTCCACTAGTTATTAGATGAAA 2481
Qy 489 ----- 489
Db 2482 ATGAAACTCCAGTATAACTGCAAGCTTGAAACTTCATTAGACTGTAAATTTTGATTATTC 2541
Qy 490 -----AlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGluGlyLeuAla 508
Db 2542 CTCAGGCGAGGCCATTGCTTACGAAAGCAATCTCTGAAGAAAGTTTGTGACTTCTTGTAG 2601
Qy 508 sAlaLysMetAsnAsnGlyTyrAspArgGluGluMetAlaArgMetValAlaCysAla 528
Db 2602 ACTCAAGTTGGAGAAAGAAATTTACGATACAACTCAGATGGCAACATGCTGCTTGTGCTG 2661
Qy 528 laAlaCysValArgHisSerAlaArgArgProArgMetSerGlnle----- 544
Db 2662 CTGCTTGCAATACGCCAATCAGCTTGGCTTCGGCTCCTAGATGAGCCAGGT-CTGAGATTTA 2720
Qy 544 ----- 544
Db 2721 GTTTAAATACATGATTGCTCCATATATCCGAAAGGAACAAGCCTAATCCATGAATACAT 2780
Qy 545 -----ValArgAlaLeuGluGlyAsnValSerLeuS 555
Db 2781 TTATATCTTGAAAAAACTTGAATAGGTAGTGTGCTCTTGAAGCGGAGGTGGCCCTGA 2840
Qy 555 erAspLeuAsnGluGlyMetArgProGlyGlnSerAsnValTyrSerSerTyrGlyGlys 575
Db 2841 GAAAGGTCGAAGAG-----ACTGGAATAGCGTGACCTATAGCTCTTCTGAAAAACC 2891
Qy 575 erThrAspTyrAspSerSerGlnTyrAsnGluAspMetLysLysPheArgLysMetAlaL 595
Db 2892 CGAATGACATC---ACACCACGGTATGGAAACAAATAAGAGAGATTC----- 2935
Qy 595 euGlyThrGlnGluTyrAsnAlaThrGlyGluTyrSerAsnProThrSerAspTyrGlyL 615
Db 2936 -----GACACAGGTTCAAGCGATGGTTACACTTCAGAAATATGAG 2975
Qy 615 euTyrProSerGlySerSerGluGlyGlnThrThrArgGluMetGluMetGlyLysI 635
Db 2976 TTAACCTTCTCAGTCGAGCAGTGAACATCAACAGGTGAATACTTAGTTTCACAGGTTCAA 3035
Qy 635 leLysA-gThrGlyGlnGlyTyrSerGlyProSerLeu 647
Db 3036 TA-----GGCAAGTTTTCACCACCAATTA 3058
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 24, 2004, 20:12:36 ; Search time 3601 Seconds
(without alignments)
5365.407 Million cell updates/sec

Title: US-10-086-464-2
Perfect score: 3453
Sequence: 1 MSSAPSGTGTGPPSPNST.....REMEMGKIKRTGGYSGPSL 647

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 segs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10086464/runat_23042004_083015_5625/app_query.fasta_1.839
-DB=EST -QFMT=fastcap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human4.0.cdi -LIST=45
-LOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10086464 @CGN 1 1 4237 @runat_23042004_083015_5625 -NCPU=6 -ICPU=3
-NO_MMAB -LARGESQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hcc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hcc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1092.5	31.6	1084	12	BP184771	BP184771 BP184771
2	1090.5	31.6	824	14	CD435142	CD435142 EL01N0355
3	1086	31.5	788	14	CF436437	CF436437 EST672782
4	1071	31.0	789	14	CF436655	CF436655 EST73000
5	1069	31.0	759	12	BM408099	BM408099 EST582426
6	1066	30.9	785	14	CF436122	CF436122 EST672467
7	1065	30.8	757	12	BG596561	BG596561 EST495239
8	1063	30.8	731	14	CB655196	CB655196 OSJNEC08F
9	1047	30.3	692	12	BG441204	BG441204 GA_Ea001
10	1030	29.8	723	14	CD839231	CD839231 RF02_114I
11	1027.5	29.8	1053	11	AY108241	AY108241 Zea mays
12	1025	29.7	1016	11	AY108243	AY108243 Zea mays
13	1018	29.5	666	12	BM358715	BM358715 GA_Ea001
14	1013	29.3	695	13	CA072174	CA072174 SCCAM100
15	983	28.5	715	13	BU099573	BU099573 WHE3309_A
16	961	27.8	770	13	BU099193	BU099193 QG21G04
17	931	27.0	611	14	CA237156	CA237156 SCWCF1500
c	931	27.0	611	14	CA237156	CA237156 SCWCF1500
19	929.5	26.9	809	14	CA765135	CA765135 AF53-Rpf
20	928	26.9	850	29	CC668754	CC668754 OGBAB66TC
c	923.5	26.7	769	14	CK283399	CK283399 EST746121
22	921	26.7	618	14	CA269355	CA269355 SCBGR7307
23	921	26.7	651	14	CF478389	CF478389 RTW3_18
24	920	26.6	673	14	CA164704	CA164704 SCBRR2312
c	919.5	26.6	938	29	CG333846	CG333846 OGGAG30TV
c	917.5	26.6	949	29	CG436431	CG436431 OGVGH77TV
27	912	26.4	679	13	BQ404121	BQ404121 GA_EG006
28	904.5	26.2	674	14	CA298046	CA298046 SCWCF1802
29	903	26.2	595	14	CF015663	CF015663 QBL8b11.x
30	903	26.2	637	13	BQ334241	BQ334241 1091014D0
31	903	26.2	645	13	BQ506869	BQ506869 EST614284
c	897.5	26.0	899	14	CK096498	CK096498 UB16CPE10
33	895.5	25.9	682	13	CA095337	CA095337 SCCCL1500
34	895	25.9	641	14	CA248316	CA248316 SCWCF1509
c	895	25.9	725	29	CG450877	CG450877 OGTAK95TH
c	893.5	25.9	803	14	CF243440	CF243440 3530_121
37	891	25.8	646	13	BQ240617	BQ240617 TAB05015B
38	891	25.8	652	13	BQ849683	BQ849683 QGB10120.
39	891	25.8	666	13	BU499754	BU499754 946178G10
40	891	25.8	691	14	CA237152	CA237152 SCWCF1500
41	889	25.7	626	13	BQ582873	BQ582873 E012278-0
42	888.5	25.7	593	10	BF176907	BF176907 EMI_4_B10
43	887	25.7	622	14	CA999610	CA999610 S345P_H09
44	887	25.7	673	13	BQ410602	BQ410602 GA_EG003
45	886.5	25.7	573	9	AV551753	AV551753 AV551753

ALIGNMENTS

RESULT 1
BP184771
LOCUS
DEFINITION BP184771 pNS rice panicle cDNA, germ cell generating stage Oryza sativa (japonica cultivar-group) cDNA, mRNA sequence.
ACCESSION BP184771
VERSION BP184771.1 GI:32948199
KEYWORDS EST.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.

```

REFERENCE
AUTHORS      1. (bases 1 to 1084)
TITLE        Moriguchi,K., Ito,Y., Yamazaki,Y. and Kurata,N.
JOURNAL      Finding of various plant nuclear proteins using yeast nuclear
COMMENT      transportation trap system - a proteomal approach
              Unpublished (2003)
              Contact: Kazuki Moriguchi
              Plant Genetics
              National Institute of Genetics
              Yata 1111, Mishima, Shizuoka 411-8540, Japan
              Tel: 81-55-981-6872
              Fax: 81-55-981-6879
              Email: kmoriguc@lab.nig.ac.jp
              cDNA clone obtained from nuclear transportation trap system
              encoding a protein similar to Oryza sativa (japonica
              cultivar-group) putative receptor protein kinase PERK1.
              Location/Qualifiers
                1..1084
                  /organism="Oryza sativa (japonica cultivar-group)"
                  /mol_type="mRNA"
                  /db_xref="taxon:39947"
                  /tissue_type="panicle"
                  /dev_stages="germ cell generating stage"
                  /clone_lib="pNS rice panicle cDNA, germ cell generating
                  stage"

FEATURES
source
ORIGIN
Alignment Scores:
Pred. No.:      4.98e-26      Length:      1084
Score:          1092.50      Matches:     212
Percent Similarity: 82.88%      Conservative: 30
Best Local Similarity: 72.60%      Mismatches: 43
Query Match:    31.64%      Indels:      7
DB:             12      Gaps:        3

US-10-086-464-2 (1-647) x BP184771 (1-1084)

Qy 360 LeuHisGlyGluGlyArgProThrMetGluTrpSerThrArgLeuLeuValAlaLeuGly 379
Db 2 TTGCACGGAAAGGCGCCCAACAAATGAGTGCGCCCAACAGACTAAAGATTGCTTTGGGA 61
Qy 380 SerAlaLysGlyLeuSerTyrLeuHisGluAspCysAsnProLysIleIleHisArgAsp 399
Db 62 GCTGCAGAAAGGTTAGCTTAATCTTCATGAAGACTGCCATCTCAAGATCATCATCGTAT 121
Qy 400 IleLysAlaSerAsnIleLeuLeuAspPheLysPheGluAlaLysValAlaAspPheGly 419
Db 122 ATTAAGGGCTCAAAACATCTCTTGATTTTAAGTTTGAATCTAAGGTTGCTGATTTTGA 181
Qy 420 LeuAlaLysIleAlaSerAspThrAsnThrHisValSerThrArgValMetGlyThrPhe 439
Db 182 CTTGCTAAGTTCACGAGTGATAATACACTCATGTTTCGCAAGAGTAAATGGCACATTTT 241
Qy 440 GlyTyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPhe 459
Db 242 GGATATCTAGCACACAGAGTACGATCTCTTGCAAGCTCACTGAGAAATCATGATGCTCTC 301
Qy 460 SerPheGlyValValLeuLeuGluLeuIleThrGlyArgArgProValAspAlaAsnAsn 479
Db 302 TCCTATGGAGTTATGCTCTTGAGTTAATAACTGCTGCGCCAGTTGTATACAGTCAA 361
Qy 480 ValTyrValAspAspSerLeuValAspTrpAlaArgProLeuLeuAsnArgAlaSerGlu 499
Db 362 ACATATATGGATAGACAGCTTGTTGACTGGGCAAGCGCTTTACTGATCGAAGCACTTGG 421
Qy 500 GlnGlyAspPheGluGlyLeuAlaAspAlaLysMetAsnAsnGlyTyrAspArgGluGlu 519
Db 422 AATGGTAACACGAGGAGTTAGTAGATCTCTGGCTTGGGAAGGATTTCAATCCCAATGAG 481
Qy 520 MetAlaArgMetValAlaCysAlaAlaAlaCysValArgHisSerAlaArgArgPro 539
Db 482 ATGGCGAGAAATGATGCTGTGTGACAGCTGCATGCGTACGCCAATTCGCTGCTGCGCCA 541
Qy 540 ArgMetSerGlnIleValArgAlaLeuGluGlyAsnValSerLeuSerAspLeuAsnGlu 559

```

```

Db 542 CGCATGAGCCAGGTGTGCGGGCTTTGGAAGGTGACGTGTCTTTGGAGGATCTTAATCAA 601
Qy 560 GlyMetArgProGlyGlnSerAsnValTyrSerSerTyrGlyGlySerThrAspTyrAsp 579
Db 602 GGTGTTCCGGCTGGTTCACAGCCGCTATTTTGGATCGTAC--AGCAGCTCTGACTATGAT 658
Qy 580 SerSerGlnTyrAsnGluAspMetLysLysPheArgLysMetAlaLeuGlyThrGlnGlu 599
Db 659 TCTGGCCAATACACAGGAGCATGAAGAAGTTCAGGAAGATGGCTTTT--ACCAACAAT 715
Qy 600 TyrAsnAlaThrGlyGlyTyrSerAsnProThrSerAspTyrGlyLeuTyrProSerGly 619
Db 716 AATGATACGACGAGCTCAATACAGCGCACCAACACGAGATATGCCAGATACCTCTGCA 775
Qy 620 SerSerSerGluGlyGlnThrThrArgGluMetGluMetGlyLysIleLysArgThrGly 639
Db 776 TCAAGCAGCGAGGGCCCAACAAACCCAGAAATCGAGACGAGCAACCAAGAGAGGGCGC 835
Qy 640 -----GlnGlyTyrSerGlyProSer 646
Db 836 TACAGTGGCTACAGCTCAGGATACAGCGGAGCCTCA 871

RESULT 2
CD435142      824 bp      mRNA      linear      EST 03-JUN-2003
LOCUS        E01N035SD03.b Endosperm_3 Zea mays cDNA, mRNA sequence.
ACCESSION    CD435142
VERSION      CD435142.1 GI:31350785
KEYWORDS     EST.
SOURCE       Zea mays
ORGANISM     Zea mays
              Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
              clade; Panicoideae; Andropogoneae; Zea.
REFERENCE    1 (bases 1 to 824)
AUTHORS      Lai,J., Dey,N., Kim,C.S., Becraft,P., Larkins,B., Linton,E. and
              Messing,J.
              Sequencing of the maize endosperm ESTs
              Unpublished (2002)
              Contact: Lai, Jinsheng
              Dr. Joachim Messing's lab
              Wakeman Institute, Rutgers University
              190 Frelinghuysen Rd., Piscataway, NJ 08854, USA
              Tel: 732-445-3801
              Fax: 732-445-5735
              Email: jlai@wakeman.rutgers.edu
              Seq primer: T3.
FEATURES      Location/Qualifiers
              1..824
                /organism="Zea mays"
                /mol_type="mRNA"
                /db_xref="taxon:4577"
                /tissue_type="Endosperm of 7-23DAP"
                /clone_lib="Endosperm_3"
                /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
                XhoI"
ORIGIN
Alignment Scores:
Pred. No.:      4.54e-26      Length:      824
Score:          1090.50      Matches:     206
Percent Similarity: 84.12%      Conservative: 27
Best Local Similarity: 74.37%      Mismatches: 41
Query Match:    31.58%      Indels:      3
DB:             14      Gaps:        2

US-10-086-464-2 (1-647) x CD435142 (1-824)

Qy 334 LeuValGlyTyrCysIleAlaGlyAlaLysArgLeuLeuValTyrGluPheValProAsn 353
Db 2 TTGGTTGGCTATTGCAITTTCTGGAGGACCTTGTGCTGTGCTATGATGATTTGTCCCAAT 61

```



```

|||||GGGTATCTGGCACCAGATATGATCCCTCTGGTAAACTAAGTAAATCAGATGCTCTTT 601
Db
460 SerPheGlyValValLeuLeuGluLeuIleThrGlyArgArgProValAspAlaAsnAsn 479
Qy
602 TCATTGGAGTCATGCTTTGGAGCTGATTACTGGAAAGCCGACCGATTGATGCAACCAA 661
Db
480 ValTyrValAspAspSerLeuValAspTrpAlaArgProLeuLeuAsnArgAlaSerGlu 499
Qy
662 ACTTATACAGATGATAGTTTGGTCGATTGGGCAAGACCATTGCTGCAGAGGCTATGGAG 721
Db
500 GlnGlyAspPheGluGlyLeuAlaAspAlaLysMetAsnAsnGlyTyrAspArgGluGlu 519
Qy
722 CATGATGATTATGATGAGCTCATTCACAGAAGGCTTAGGAGATAAATCAACCATGACGAA 781
Db
520 MetAla 521
Qy
782 ATGGCA 787
Db

RESULT 4
CF436655
LOCUS EST673000 normalized cDNA library of onion Allium cepa linear EST 04-SEP-2003
DEFINITION ACACK26 mRNA sequence.
ACCESSION CF436655
VERSION CF436655.1 GI:34459345
KEYWORDS EST.
SOURCE Allium cepa (onion)
ORGANISM Allium cepa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Alliaceae;
Allium.
REFERENCE 1 (bases 1 to 789)
Havely,M.J., Cheung,F., Van Aken,S., Utterback,T. and Town,C.D.
Expressed Sequence Tags from a normalized library of mixed onion
tissues (Allium cepa)
JOURNAL Unpublished (2003)
COMMENT Contact: Havely MJ
Department of Horticulture
USDA-ARS and University of Wisconsin
1575 Linden Drive, Madison, WI 53706, USA
Tel: 608-262-1830
Fax: 608-262-4743
Email: mjhavely@facstaff.wisc.edu
TIGR sequence name ACACK26TFR. For more information:
http://havelylab.hort.wisc.edu
Seq primer: CAG GAA ACA GCT ATG ACC.
FEATURES
source
Location/Qualifiers
1..789
/organism="Allium cepa"
/mol_type="mRNA"
/cultivar="Red Creole (bulbs), unknown (callus), Ebano &
Texas Legend (roots)"
/db_xref="taxon:4679"
/clone="ACACK26"
/tissue_type="Callus, roots, and young bulbs"
/clone_lib="normalized cDNA library of onion"
/note="Vector: pCMVSPORT6.1-ccdb (Invitrogen); Site 1:
EcoRV (5'); Site 2: NotI (3'); Equal molar amounts of mRNA
from callus, roots, and young bulbs were combined to
synthesize the library. Normalization was performed by
low-copy transcripts was performed by proprietary
techniques of Invitrogen."

ORIGIN
Alignment Scores:
Pred. No.: 1,82e-25 Length: 789
Score: 1071.00 Matches: 203
Percent Similarity: 86.64% Conservative: 24
Best Local Similarity: 77.48% Mismatches: 35
Query Match: 31.02% Indels: 0
DB: 14 Gaps: 0

```

```

US-10-086-464-2 (1-647) x CF436655 (1-789)
Qy 260 LysSerThrPheThrTyrGluGluLeuAlaArgAlaThrAsnGlyPheSerGluAlaAsn 279
Db 2 AAGAGCATTTCACATATGAAGAATTGGCGATAGTACAAATGGGTTCGCACTAATAAT 61
Qy 280 LeuLeuGlyGlnGlyGlyPheGlyTyrValHisLysGlyValLeuProSerGlyValGlu 299
Db 62 CTTCTTGGGAAGGTGGATTGGATATGTGCACAAAGGAGTACTTCCAAACGGTAAGAA 121
Qy 300 ValAlaValLysGlnLeuLysValGlySerGlyGlnGlyGluArgGluPheGlnAlaGlu 319
Db 122 GTGGCTATCAAGCAGTTGAAAGCCGGAAGCGCAAGGGGAGCGTGAGTTTCAAGCAGAG 181
Qy 320 ValGluIleSerArgValHisArgHisLeuValSerLeuValGlyTyrCysIle 339
Db 182 GTTGAATAATTATCAGTCGAGTGCATCATGAGCATTTGGTTCTTTAGTTGGCTATTGCATT 241
Qy 340 AlaGlyAlaLysArgLeuLeuValTyrGluPheValProAsnAsnLeuGluLeuHis 359
Db 242 TCAGGAGATCATGATTGCTTGTCTATGAATATGTTCTTAATAAAACCCCTGAGTTCAT 301
Qy 360 LeuHisGlyGluGlyArgProThrMetGluTrpSerThrArgLeuLysIleAlaLeuGly 379
Db 302 TTGCATGGAAAGATGTTCCACCTATGATGGCAACGGGTTAAAAATTTGCTTTGGGT 361
Qy 380 SerAlaLysGlyLeuSerTyrLeuHisGluAspCysAsnProLysIleIleHisArgAsp 399
Db 362 TCTGCCAAGGGTTGGCATATCTCCATGAAGATTGTCCATCCCAAAATTAATTTCATCGTAT 421
Qy 400 IleLysAlaSerAsnIleLeuLysAspPheLysGluAlaLysValAlaAspPheGly 419
Db 422 ATTAAGACAGCAAAATATTTCTTTGAGATGAACATTTGAGGCTAAGGTTGAGATTTTGGC 481
Qy 420 LeuAlaLysIleAlaSerAspThrAsnThrHisValSerThrArgValMetGlyThrPhe 439
Db 482 CTTGCAAGTTCACTGTTGAAATAACACCCATGTTTCTACACGAGTCATGGGAACATTT 541
Qy 440 GlyTyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPhe 459
Db 542 GGGTATCTGGCACCAGAAATATGCATCTCTGGTAAACTAAGTGTATAAATCAGATGCTCTTT 601
Qy 460 SerPheGlyValValLeuLeuGluLeuIleThrGlyArgArgProValAspAlaAsnAsn 479
Db 602 TCATTTGGAGTCATGCTTTTGGAGCTGATTACTGGAAGCGACCATGTTGATGCAACCCAA 661
Qy 480 ValTyrValAspAspSerLeuValAspTrpAlaArgProLeuLeuAsnArgAlaSerGlu 499
Db 662 ACTTATACAGATGATAGTTTGGTCGATTGGCAAGACCATTTCTGCAGAGGCTATGGAG 721
Qy 500 GlnGlyAspPheGluGlyLeuAlaAspAlaLysMetAsnAsnGlyTyrAspArgGluGlu 519
Db 722 CATGATGATTATGATGAGCTCATTCACAGAAGGCTTAGGAGATAAATACCCCATGACGA 781
Qy 520 MetAla 521
Db 782 ATGGCA 787

RESULT 5
BM408099
LOCUS EST582426 potato roots Solanum tuberosum cDNA clone cPRO33021 5'
DEFINITION end, mRNA sequence.
ACCESSION BM408099
VERSION BM408099.1 GI:18259729
KEYWORDS EST.
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum.
REFERENCE 1 (bases 1 to 759)
AUTHORS van der Hoeven,R., Sun,H., Karamycheva,S.A., Tsai,J., Van Aken,S.,

```

Utterback, T., Chiemingo, A., Bougri, O., Buell, C.R., Romning, C., Tanksley, S. and Baker, B.
 Generation of ESTs from potato roots
 Unpublished (2001)
 Contact: Robin Buell
 The Institute for Genomic Research
 9712 Medical Center Dr, Rockville, MD 20850, USA
 Email: potato-array@tigr.org
 This clone can be obtained from the University of Arizona Genomics Institute. Orders can be made through URL:
<http://genome.arizona.edu/orders/>
 Seq primer: T3.

FEATURES

Location/Qualifiers
 1. 759
 /organism="Solanum tuberosum"
 /mol_type="mRNA"
 /cultivar="Kennebec"
 /db_xref="taxon:4113"
 /clone="CPRO33021"
 /tissue_type="roots"
 /dev_stage="in vitro grown stem cuttings"
 /lab_host="SOLR"
 /clone_lib="potato roots"
 /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2: XhoI; supplier: Cornell University, Tanksley lab; sequencing: The Institute for Genomic Research. Roots were isolated from in vitro grown stem cuttings on CM medium. Roots were isolated two weeks after placing the stem cuttings from in vitro grown plants on medium."

ORIGIN

Alignment Scores:
 Pred. No.: 2,04e-25 Length: 759
 Score: 1069.00 Matches: 200
 Percent Similarity: 89.29% Conservative: 25
 Best Local Similarity: 79.37% Mismatches: 27
 Query Match: 30.96% Indels: 0
 DB: 12 Gaps: 0

US-10-086-464-2 (1-647) x BM408099 (1-759)

Qy 312 GlyGluArgGluPheGlnAlaGluValGluLeuSerArgValHisArgHisLeu 331
 Db 2 GGGGAACGTGAATTCAGCGCGAAGTTGAGATTATTAGCCGAGTACATCAAGCATCTT 61
 Qy 332 ValSerLeuValGlyTyrCysIleAlaGlyAlaLysArgLeuLeuValTyrGluPheVal 351
 Db 62 GTGTCCTCTTGTGGATATCGCATTTACTGGGGCTCAGAGACTGCTTGTATTGAGTTTGT 121
 Qy 352 ProAsnAsnLeuLeuGluLeuHisGlyGluGlyArgProThrMetGluTrpSer 371
 Db 122 CCAAACTACTTTGGAAATTCATTTACAGGAAGGAGCCCTCTTTGGATTGGCCA 181
 Qy 372 ThrArgLeuLysIleAlaLeuGlySerAlaLysGlyLeuSerTyrLeuHisGluAspCys 391
 Db 182 ATACGGCTAAAGATTGCTCTAGGTCAGCTAAAGGACTGGCATATCTGCATGAGACTGC 241
 Qy 392 AsnProLysIleIleHisArgAspIleLysAlaSerAsnIleLeuIleAspPheLysPhe 411
 Db 242 CAACCGAAATCAATTCACCGGTGATATCAAGCGAGCTAATATATCTATCGACTTTAATTT 301
 Qy 412 GluAlaLysValAlaAspPheGlyLeuAlaLysIleAlaSerAspThrAsnThrHisVal 431
 Db 302 GAGGCTAAGGTGCTGATTTTGGACTTGGCAAGCTAACTTCTGATGTTAATATCTCATGTC 361
 Qy 432 SerThrArgValMetGlyThrPheGlyTyrLeuAlaProGluTyrAlaAlaSerGlyLys 451
 Db 362 TCCACGAGGTGATGGGAACCTTTGGGTATTGGCTCCAGATATGCTTCTCTGGAAG 421
 Qy 452 LeuThrGluLysSerAspValPheSerPheGlyValValLeuLeuGluLeuIleThrGly 471
 Db 422 CTTACAGACAAGTCAGACGTAATTCCTCTTGGTGAATGCTTCTTGATGTTGATAACTGGA 481

Qy 472 ArgArgProValAspAlaAsnAsnValTyrValAspAspSerLeuValAspTrpAlaArg 491
 Db 482 CGTCGGCCTGTTGACTCTACTCAATCATCATCAAGATAGTTTGGTGGACTGGCAGCT 541
 Qy 492 ProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGluGlyLeuAlaAspAlaLysMet 511
 Db 542 CCATTACTCACAGAGCTTTAGAGAGATGAAGAAGTTTGATACCCCTTGTGATCGTCGGTA 601
 Qy 512 AsnAsnGlyTyrAspArgGluGluMetAlaArgMetValAlaCysAlaAlaCysVal 531
 Db 602 GAAATGATTATTAACCATATAGATGCTCGCATGTTGCTTGTGCTGCTTGTGTG 661
 Qy 532 ArgHisSerAlaArgArgProArgMetSerGlnIleValArgAlaLeuGluGlyAsn 551
 Db 662 CGTCATTTCAGCAAGCGTAGACCAAGTACACAGGTTCTCCGAGCCTTGGAAGGAGAT 721
 Qy 552 ValSerLeuSerAspLeuAsnGluGlyMetArgPro 563
 Db 722 GTCTCATTTACAGACCTTTACGAAGGAGTTAAACCT 757

RESULT 6

LOCUS CF436122 785 bp mRNA linear EST 04-SEP-2003
 DEFINITION EST672467 normalized cDNA library of onion Allium cepa cDNA clone ACACJ26, mRNA sequence.

ACCESSION CF436122

VERSION CF436122.1 GI:34458812

KEYWORDS EST.

SOURCE Allium cepa (onion)

ORGANISM Allium cepa

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Alliaceae; Allium.

REFERENCE 1 (Bases 1 to 785)

AUTHORS Havey, M.J., Cheung, F., Van Aken, S., Utterback, T. and Town, C.D.

TITLE Expressed Sequence Tags from a normalized library of mixed onion tissues (Allium cepa)

JOURNAL Unpublished (2003)

COMMENT Contact: Havey MJ

Department of Horticulture

USDA-ARS and University of Wisconsin

1575 Linden Drive, Madison, WI 53706, USA

Tel: 608-262-1830

Fax: 608-262-4743

Email: mjhavey@facstaff.wisc.edu

TIGR sequence name ACACJ26TR. For more information:

<http://haveylab.hort.wisc.edu>

Seq primer: CAG GAA ACA GCT ATG ACC.

FEATURES

Location/Qualifiers

1. 785

/organism="Allium cepa"

/mol_type="mRNA"

/cultivar="Red Creole (bulbs), unknown (callus), Ebano &

Texas Legend (roots)"

/db_xref="taxon:4679"

/clone="ACACJ26"

/tissue_type="Callus, roots, and young bulbs"

/clone_lib="normalized cDNA library of onion"

/notes="Vector: pCMVSPORT6.1-ccdb (Invitrogen); Site 1: EcoRV (5'); Site 2: NotI (3'); Equal molar amounts of mRNA from callus, roots, and young bulbs were combined to synthesize the library. Normalization to enrich for low-copy transcripts was performed by proprietary techniques of Invitrogen."

ORIGIN

Alignment Scores:

Pred. No.: 2,61e-25 Length: 785

Score: 1066.00 Matches: 204

Percent Similarity: 86.64% Conservative: 23

Best Local Similarity: 77.86% Mismatches: 34

Query Match: 30.87% Indels: 1

DB: 14 Gaps: 0

```
US-10-086-464-2 (1-647) x CF436122 (1-785)

Qy 260 LysSerThrPheThrTyrGluGluLeuAlaArgAlaThrAsnGlyPheSerGluAlaAsn 279
Db 2 AAGAGCACTTTACATATGAAGATTGGCGATAGCTACAAATGGTTTCCGACTATAAT 61

Qy 280 LeuLeuGlyGlnGlyGlyPheGlyTyrValHisGlyValLeuProSerGlyLysGlu 299
Db 62 CTTCCTGGCCAAAGTGGATTGGATATGTGCACAAAGGAGTACTTCCAAACGGTAAAGAA 121

Qy 300 ValAlaValLysGlnLeuLysValGlySerGlyGlnGlyGluArgGluPheGlnAlaGlu 319
Db 122 GTGGCTATCAAGCAGTTGAAGCCGGAAGCGGCAAGGGGAGCGTGAGTTCAAGCAGAG 181

Qy 320 ValGluIleIleSerArgValHisHisArgHisLeuValSerLeuValGlyTyrCysIle 339
Db 182 GTTGAATATACAGTCGAGTCATCATAGGCATTTGGTTCTTTAGTTGGCTATTGCATT 241

Qy 340 AlaGlyAlaLysArgLeuLeuValTyrGluPheValProAsnAsnLeuGluLeuHis 359
Db 242 TCAGGAGATCATAGATTCTTGTCTATGATATATGTTCTTAATAAAACCCCTTGAGTTCCAT 301

Qy 360 LeuHisGlyGluGlyArgProThrMetGluTrpSerThrArgLeuLysIleAlaLeuGly 379
Db 302 TTGCATGGAAAGATGTTCCACTATGATTGGCCACGCGTTAAATAATTCGTTTGGGT 361

Qy 380 SerAlaLysGlyLeuSerTyrLeuHisGluAspCysAsnProLysIleIleHisArgAsp 399
Db 362 TCTGCCAAGGGTTGGCATATCTCCATGAAGATTGTCATCCCAAAATTTATTCATCGTAT 421

Qy 400 IleLysAlaSerAsnIleLeuIleAspPheLysPheGluAlaLysValAlaAspPheGly 419
Db 422 ATTAAGCAGCAAAATATTCTTTCAGATGAACTTTGAGGCTAAGGTTGCGAGATTTGGC 481

Qy 420 LeuAlaLysIleAlaSerAspThrAsnThrHisValSerThrArgValMetGlyThrPhe 439
Db 482 CTTCGAAAGTTCACGTGGTGAATAAACACCCATGTTTCTACACGAGTCATGGGAACATTT 541

Qy 440 GlyTyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPhe 459
Db 542 GGGTATCTGGCCAGCATATGATCTCTCTGTAAACTAAGTATAATCATAGATGCTTTT 601

Qy 460 SerPheGlyValValLeuGluLeuIleThrGlyArgArgProValAspAlaAsnAsn 479
Db 602 TCATTGGAGTCATGCTTTTGGAGCTGATTACTTGAAAGCGACCGAGTTGATGCAACCCAA 661

Qy 480 ValTyrValAspAspSer-LeuValAspTrpAlaArgProLeuLeuAsnArgAlaSerG1 499
Db 662 ACTTATACAGATGATAGATTTTGGTCGATTGGCAGACCATGCTGCGAAGGCTATGGA 721

Qy 499 uGlnGlyAspPheGluGlyLeuAlaAspAlaLysMetAsnAsnGlyTyrAspArgGluG1 519
Db 722 GCATGATGATTATGATGAGCTCAITTGACAGAAAGCGCTAGGAGATAATAACAACCATGACGA 781

Qy 519 uMet 520
Db 782 AATG 785

RESULT 7
LOCUS BG596561
DEFINITION EST495239 cSTS Solanum tuberosum cDNA clone cSTS15A23 5' sequence, mRNA sequence.
ACCESSION BG596561
VERSION BG596561.1 GI:13614701
KEYWORDS EST.
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
TAXID 3845
XREF: Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots;
asterids; lamiales; Solanales; Solanaceae; Solanum.
REFERENCE 1 (bases 1 to 757)

AUTHORS van der Hoeven,R., Bezzerides,J., Sun,H., Cho,J., Chiemingo,A.,
Bougri,O., Buell,C.R., Ronning,C., Tanksley,S. and Baker,B.
TITLE Generations of ESTs from sprouting potato eyes
JOURNAL Unpublished (2000)
COMMENT Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
This clone can be obtained from the University of Arizona Genomics
Institute. Orders can be made through URL:
http://genome.arizona.edu/orders/
Seq primer: M13P-R.
FEATURES
source
1..757
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="cSTS15A23"
/tissue_type="sprouting eyes from tubers"
/dev_stage="12-14 weeks post harvest"
/lab_host="SOLR"
/clone_lib="cSTS"
/notes="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; Various sizes of sprouting eyes (2mm to 15mm) were
taken from tubers. The tubers were incubated at 26C in the
dark for 2-3 weeks prior to sprouting. The eyes were
frozen in liquid nitrogen immediately upon removal from
tubers."
ORIGIN
Alignment Scores: 2.72e-25 Length: 757
Pred. NO.: 301 AlaValLysGlnLeuLysValGlySerGlyGlnGlyGluArgGluPheGlnAlaGluVal 320
Score: 1065.00 Matches: 200
Percent Similarity: 89.16% Conservative: 22
Best Local Similarity: 80.32% Mismatches: 27
Query Match: 30.84% Indels: 0
DB: 12 Gaps: 0
US-10-086-464-2 (1-647) x BG596561 (1-757)
Qy 301 AlaValLysGlnLeuLysValGlySerGlyGlnGlyGluArgGluPheGlnAlaGluVal 320
Db 1 GCAGTTAAACAGCTTAAGCGTGGAAAGTGGACAGAGGGAACGTGAATTTCCAGCGGAAGTT 60
Qy 321 GluIleIleSerArgValHisHisArgHisLeuValSerLeuValGlyTyrCysIleAla 340
Db 61 GAGATTATTAGCCGAGTACATCACAGCATCTTGTGTCTCTTGTGGATACTGCGATTACT 120
Qy 341 GlyAlaLysArgLeuLeuValTyrGluPheValProAsnAsnAsnLeuGluLeuHisLeu 360
Db 121 GGGGCTCAGAGACTGCTTGTATTATGAGTTTGTCCAAACAATACTTTTGGAAATTTTCATTTA 180
Qy 361 HisGlyGluGlyArgProThrMetGluTrpSerThrArgLeuLysIleAlaLeuGlySer 380
Db 181 CACGAAAGGGAAGCGCTCTCTTGGATTGGCCAATACGGCTAAAGATTGCTCTAGGGTCA 240
Qy 381 AlalysGlyLeuSerTyrLeuHisGluAspCysAsnProLysIleIleHisArgAspIle 400
Db 241 GCTAAGGAGCTGGCATATCTGCATGAAGACTGCCAACCGAAATCAATTCACCGTGATATC 300
Qy 401 LysAlaSerAsnIleLeuIleAspPheLysPheGluAlaLysValAlaAspPheGlyLeu 420
Db 301 AAGGCAGCTAATATACCTATTCGACATTTAATTTTGGAGCTAAGGTTGCTGATTTTGGACTT 360
Qy 421 AlalysIleAlaSerAspThrAsnThrHisValSerThrArgValMetGlyThrPheGly 440
Db 361 GCCAAGCTAACTTCTGATGTTAATATCTCATGTCCTCCACGAGAGTGATGGGAACTTTTGGG 420
Qy 441 TyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPheSer 460
Db 421 TATTGGCTCCAGAAATATGCTTCCTCTGGAAAGCTTACAGACAAGTACAGACGATTTCTCC 480
```



```
JOURNAL
COMMENT
of the cotton fiber
Unpublished (2000)
Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Seq primer: TAATACGACTCACTATAGGG
High quality sequence stop: 690.

FEATURES
    source
        1..692
            /organism="Gossypium arboreum"
            /mol_type="mRNA"
            /strain="AKA"
            /cultivar="8400"
            /db_xref="taxon:29729"
            /clone="GA_E40012C15f"
            /tissue_type="Fibers isolated from bolls harvested 7-10 dpa"
            /lab_host="E. coli"
            /clone_lib="Gossypium arboreum 7-10 dpa fiber library"
            /note="Vector: pBK-CMV; Site_1: EcoRI; Site_2: XhoI"

ORIGIN
Alignment Scores:
Pred. No.:      9.38e-25      Length:      692
Score:          1047.00      Matches:    198
Percent Similarity: 94.71%      Conservative: 17
Best Local Similarity: 87.22%      Mismatches: 12
Query Match:    30.32%      Indels:     0
DB:              12          Gaps:         0

US-10-086-464-2 (1-647) x BG441204 (1-692)

Qy 247 LeuProProSerProGlyLeuValLeuGlyPheSerLysSerThrPheThrTy-Glu 266
Db 11 CTTCCGCTCGGTCCACCTGGATTTCTCTTGGTTCGAAAGACATTTTAGCTATGAA 70
Qy 267 GluLeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyPhe 286
Db 71 GAATAGCGAGACGACGATGGCTTCCTCGGAGTTAACCTTCCTGGACAGGTGTTT 130
Qy 287 GlyTyrValHisLysGlyValLeuProSerGlyLysGluValAlaValLysGlnLys 306
Db 131 GGGTACGTACACAAAGGATTCCTCCTAATGGGAGGAGTAGCAGTAAGCAACTCAAG 190
Qy 307 ValGlySerGlyGlnGlyGluArgGluPheGlnAlaGluValGluLeuLeuSerArgVal 326
Db 191 GCTGGAAGTGGCGAAGCGAGAGAGATTTCCAGGCTGAAGTTGAGATCATTTAGCCGCTC 250
Qy 327 HisHisArgHisLeuValSerLeuValGlyTyrCysIleAlaGlyAlaLysArgLeuLeu 346
Db 251 CATCACAAACATCTGCTCTCATTTGGTCGATCTGTAATTCGGACAAATGAATGCTT 310
Qy 347 ValTyrGluPheValProAsnAsnLeuGluLeuHisLeuHisGlyGluGlyArgPro 366
Db 311 GTTTATGAGTTGTTTCCAAACAACACCTTGGAGTTTCACCTTGCAATGGAGGGCGACTG 370
Qy 367 ThrMetGluTyrSerThrArgLeuLysIleAlaLeuGlySerAlaLysGlyLeuSerTyr 386
Db 371 ACCATGAGTTGGCCGACGAAGGATGAATAATGCTTTAGATCTGCAAAAGGACTGGCATAT 430
Qy 387 LeuHisGluAspCysAsnProLysIleIleHisArgAspIleLysAlaSerAsnIleLeu 406
Db 431 CTTTCATGAAGATTGTCATCCTTAAGATCATTCACCGTGATATTAGGCCGCTAATATCTG 490
Qy 407 IleAspPheLysPheGluAlaLysValAlaAspPheGlyLeuAlaLysIleAlaSerAsp 426
Db 491 TTGGAATTTCAAGTTTGAAGCAAGGTTGCTGATTTTGGACTAGCGAAATTTGCTTCGAT 550
Qy 427 ThrAsnThrHisValSerThrArgValMetGlyThrPheGlyTyrLeuAlaProGluTyr 446

Db 551 GTCACACGACGCTCTCCACCAGGTCATGGGTACTTTCGGTATTTAGCCCTGAGTAT 610
Qy 447 AlaAlaSerGlyLysLeuThrGluLysSerAspValPheSerPheGlyValValLeuLeu 466
Db 611 GCTTCAAGTGGAAAGCTCACTATAAATCAGATGTTTCTCTCCTTCGGGGTCATGCTTTT 670
Qy 467 GluLeuIleThrGlyArgArg 473
Db 671 GAGTTGATTACCGGTCACAGA 691

RESULT 10
LOCUS CD839231
DEFINITION CD839231 723 bp mRNA linear EST 10-JUL-2003
sequence.
ACCESSION CD839231
VERSION CD839231.1 GI:32521171
KEYWORDS EST.
SOURCE Brassica napus (rape)
ORGANISM Brassica napus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE
1 (bases 1 to 723)
Genoplaente.
AUTHORS Genoplaente, a major partnership french program in plant genomics
TITLE Unpublished (2003)
JOURNAL Contact: Genoplaente
COMMENT Genoplaente
93, rue Henri Rochefort 91025 EVRY CEDEX France
Tel: 33 1 69 47 54 00
Fax: 33 1 69 47 54 10
This sequence has been generated in the framework of the french
plant genomics programme 'Genoplaente' (http://www.genoplaente.com
and http://genoplaente-info.infobiogen.fr).

FEATURES
    source
        1..723
            /organism="Brassica napus"
            /mol_type="mRNA"
            /cultivar="samourai (restored line)"
            /db_xref="taxon:3708"
            /clone="RFO2114104"
            /tissue_type="anthers"
            /clone_lib="RFO2"

ORIGIN
Alignment Scores:
Pred. No.:      3.37e-24      Length:      723
Score:          1030.00      Matches:    199
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    29.83%      Indels:     0
DB:              14          Gaps:         0

US-10-086-464-2 (1-647) x CD839231 (1-723)

Qy 449 SerGlyLysLeuThrGluLysSerAspValPheSerPheGlyValValLeuLeuGluLeu 468
Db 2 AGCGAAAGCTCACCGAAGTCTGACGTTTCTCAITTTGGCGTTGTGCTTTTGGAGCTC 61
Qy 469 IleThrGlyArgArgProValAspAlaAsnAsnValTyrValAspAspSerLeuValAsp 488
Db 62 ATTAACCGGGCGTCGACCGCTTGATGCCAACAAATGCTATGTAGATGACAGCTTAGTTGAC 121
Qy 489 TrpAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGlyLeuAlaAsp 508
Db 122 TGGGCACGACCATGCTTAACCGAGCATCTGAGCAAGGAGACTTTTGGAGGTTTACGTGAT 181
Qy 509 AlaLysMetAsnAsnGlyTyrAspArgGluMetAlaArgMetValAlaCysAlaAla 528
Db 182 GCAAGATGAATAATGGGTATGACAGAGAGAGATGGCTCGCATGTTGCTTGTGCTGCG 241
```

Qy 529 AlaCysValArgHisSerAlaArgArgProArgMetSerGlnIleValArgAlaLeu 548
 Db 242 GCTTGTGTCGCCATTACCTCGCGCAGACCTCGCATGAGCCAGATTGTGCGTGGTTA 301
 Qy 549 GluGlyIleValSerLeuSerAspLeuAsnGluGlyMetArgProGlyGlnSerAsnVal 568
 Db 302 GAAGGAAATGTATCACTGTCTAGATCTTAACGAAGGGATGAGACCGAGTCAAGCAATGTA 361
 Qy 569 TyrSerSerTyrGlyGlySerThrAspTyrAspSerSerGlnTyrAsnGluAspMetLys 588
 Db 362 TACAGCTCATACGAGGAGCAGCAGATATGACTCGAGCCAGTACAAATGAAGACATGAAG 421
 Qy 589 LysPheArgLysMetAlaLeuGlyThrGlnGluTyrAsnAlaThrGlyGluTyrSerAsn 608
 Db 422 AAGTTTAGGAAATGGCTCTTGGAACTCAAGAGTACAAACGCCACGGGTGAGTACAGTAAT 481
 Qy 609 ProThrSerAspTyrGlyLeuTyrProSerGlySerSerGluGlyGlnThrThrArg 628
 Db 482 CCGACCAAGCGACTATGAGCTGTACCCGCTCTGTTCAGAGCAGGAGGCGCAACACACGC 541
 Qy 629 GluMetGluMetGlyLysIleLysArgThrGlyGlnGlyTyrSerGlyProSerLeu 647
 Db 542 GAAATGGAGATGGGGAAGATTAAAGAAACCGGTGAGGTTATAGTGGACCTTCTCTT 598

RESULT 11

AY108241
 LOCUS AY108241 1053 bp mRNA linear HTC 16-OCT-2002
 DEFINITION Zea mays PC0134818 mRNA sequence.
 ACCESSION AY108241
 VERSION AY108241.1 GI:21211319
 KEYWORDS HTC.

SOURCE

ORGANISM Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 1053)
 Hainey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whitsitt, M.S., Arthur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V.

Maize Mapping Project/DuPont Consensus Sequences for Design of Overgo Probes

Unpublished (2002)
 2 (bases 1 to 1053)
 Coe, E.H.

Direct Submission

Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA

If you are interested in getting corresponding physical clones, these are publicly available from ZmDB and may be found by BLAST searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR, www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the maize cDNA sequences is either Virginia Walbot, Stanford or Pat Schnable, Iowa State, then clones may be requested from ZmDB: www.zmdb.iastate.edu.

FEATURES

source

1. 1053

/organism="Zea mays"

/mol_type="mRNA"

/db_xref="taxon:4577"

/clone_lib="Maize Mapping Project/DuPont Consensus Library"

/note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed DuPont contigs; this resource was assembled by DuPont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"

ORIGIN

Alignment Scores: 5.61e-24 Length: 1053

Pred. No.: 1027.50 Matches: 203

Score:

Percent Similarity: 79.93% Conservative: 28
 Best Local Similarity: 70.24% Mismatches: 46
 Query Match: 29.76% Indels: 12
 DB: 11 Gaps: 4

US-10-086-464-2 (1-647) x AY108241 (1-1053)

Qy 365 ArgProThrMetGluTyrSerThrArgLeuLysIle-AlaLeuGlySerAlaLysGlyLe 384
 Db 6 CGTCCG-----GAGTGGCTCTAGATTAAAGATCATGTTGGGTGCTGCAAGGGTTT 59
 Qy 384 uSerTyrLeuHisGluAspCysAsnProLysIleIleHisArgAspIleLysAlaSerAs 404
 Db 60 AGCTTATCTTCATGAAGACTGCCATCCAAAGATCATCCATCGTGCATATAAGGCATCTAA 119
 Qy 404 nIleLeuIleAspPheLysPheGluAlaLysValAlaAspPheGlyLeuAlaLysIleAl 424
 Db 120 CATCTCTCTTGCATTCCTCAATTTGAAGCTAAGGTTCGTGATTTGGACTTGCATAAGTTT 179
 Qy 424 aSerAspThrAsnThrHisValSerThrArgValMetGlyThrPheGlyTyrLeuAlaPr 444
 Db 180 TACTGATAACAACACCCATGTTTCGACACAGATGTAATGGGCACCTTTGGGTATTTGGCACC 239
 Qy 444 oGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPheSerPheGlyValVa 464
 Db 240 TGAGTATGCATCTCTCGCAAGCTAACAGAAAAATCGATGATTTTCTTCGGAGTCAT 299
 Qy 464 lLeuLeuGluLeuIleThrGlyArgArgProValAspAlaAsnValTyrValAspAs 484
 Db 300 GCTCTTTGAGCTTATTACTGGGGCGGACCATGTTGACACCAACCCCAACATATATGGATGA 359
 Qy 484 pSerLeuValAspTyrAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGl 504
 Db 360 CAGCTTGTTGACTGGCAAGGCCATTTACTGATCGGAGCACTTGAGATGTTGATATGA 419
 Qy 504 uGlyLeuAlaAspAlaLysMetAsnAsnGlyTyrAspArgGluGluMetAlaArgMetVa 524
 Db 420 TGCTTTAGTCGATCCTCGCTCGGGAAGGACTTCAATCCTTAATGAGATGGCAAGATGAT 479
 Qy 524 lAlaCysAlaAlaAlaCysValArgHisSerAlaArgArgProArgMetSerGlnI 544
 Db 480 AGCCTGTGCAGCTGCATGTGTACGCCATTTCTGCACGTCGTGCGCCACGATAGTACGTAG 539
 Qy 544 eValArgAlaLeuGluGlyAsnValSerLeuSerAspLeuAsnGluGlyMetArgProGl 564
 Db 540 CGTTGGGCTTTGAGGGCCAAATGTCTTTGGAGGACCTTAATGAAGGTGTTGGGCTGG 599
 Qy 564 yGlnSerAsnValTyrSerTyrGlyGlySerThrAspTyrAspSerSerGlnTyrAs 584
 Db 600 CCATAGCCGGCTTCTTTGGGTGCATAC---AGCAGCTCCGATTACGATTTCTGCGCCAGTACAA 656
 Qy 584 nGluAspMetLysLysPheArgLysMetAlaLeuGlyThrGlnGluTyrAsnAlaThrGl 604
 Db 657 CGAGGACATGAAGAAGTTCAAGAAGATGGCATTCATCAACAAC---TATACACGAG 710
 Qy 604 yGluTyrSerAsnProThrSerAspTyrGlyLeuTyrProSerGlySerSerGluGl 624
 Db 711 CCAATACAGCCGCCCAACCACTGATGATGACAGATACCTCTGTCATCAAGACGAGGG 770
 Qy 624 yGlnThrThrArgGluMetGluMetGlyLysIleLysArgThrGly----- 639
 Db 771 CCACCAGACCGCAAGAGATGGAGTCGGGTGCAATGAAGAAGGGTGTACAGTGTGGCTA 830
 Qy 640 ----GlnGlyTyrSerGlyProSer 646
 Db 831 CAGCTCAGGATACAGCGGAGCCTCG 855

RESULT 12

AY108243
 LOCUS AY108243 1016 bp mRNA linear HTC 16-OCT-2002
 DEFINITION Zea mays PC0134814 mRNA sequence.
 ACCESSION AY108243
 VERSION AY108243.1 GI:21211321


```

KEYWORDS      HTC.
SOURCE         Zea mays
ORGANISM       Zea mays
REFERENCE      1 (bases 1 to 1016)
AUTHORS        Hayney,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
                Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
TITLE          Maize Mapping Project/DuPont Consensus Sequences for Design of
                Overgo Probes
JOURNAL        Unpublished (2002)
AUTHORS        Coe,E.H.
REFERENCE      2 (bases 1 to 1016)
AUTHORS        Direct Submission
JOURNAL        Submitted (25-APR-2002) Maize Mapping Project, University of
                Missouri, Columbia, MO 65211, USA
COMMENT        If you are interested in getting corresponding physical clones,
                these are publicly available from ZmDB and may be found by BLAST
                searching at MSU, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
                www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
                maize cDNA sequences is either Virginia Walbot, Stanford or Pat
                Schnable, Iowa State, then clones may be requested from ZmDB:
                www.zmdb.iastate.edu.
FEATURES       Location/Qualifiers
                1..1016
                /organism="Zea mays"
                /mol_type="mRNA"
                /db_xref="MaizeDB:637888"
                /db_xref="taxon:4577"
                /clone_lib="Maize Mapping Project/DuPont Cornsensus
                Library"
                /note="this sequence is part of a project of EST
                assemblies resulting from the application of public
                contigs to seed DuPont contigs; this resource was
                assembled by DuPont as part of a collaboration for the
                overgo addressing of BACs in conjunction with the Maize
                Mapping Project"
ORIGIN
Alignment Scores:
Pred. No.:      6 53e-24      Length:      1016
Score:          1025.00      Matches:    200
Percent Similarity: 82.14%      Conservaive: 30
Best Local Similarity: 71.43%      Mismatches: 42
Query Match:    29.68%      Indels:     8
DB:             11          Gaps:         4

US-10-086-464-2 (1-647) x AY108243 (1-1016)

Qy 373 ArgLeuYsIlleAlaLeuGlySerAlaLysGlyLeuSerTyrLeuHisGluAspCysAsn 392
Db 3 AGATTAAAGATCGCTCTGGTGCTGCCAAGGGTTAGCTTATCTTCATGAAGACTGCCAT 62
Qy 393 ProYsIlleIleHisArgAspIleYsAlaSerAsnIleLeuIleAspPheLysPheGlu 412
Db 63 CCAAAGATCATCCACCGCGCATTAAGGCATCTAACATCTCTTGACTTCATAATTTGAA 122
Qy 413 AlalysValAlaAspPheGlyLeuAlalysIleAlaSerAspThrAsnThrHisValSer 432
Db 123 GCTATGTTGCTGACTTTGGACTTGCAAGTTCTACTACTGATGATAACAACCCATGTGTCA 182
Qy 433 ThrArgValMetGlyThrPheGlyTyrLeuAlaProGluTyrAlaAlaSerGlyLysLeu 452
Db 183 ACAAGAGTAATGGGCACCTTTGGGTATTGGCACCAGTAGTATGCAGCATCTGGCAAGCTC 242
Qy 453 ThrGluLysSerAspValPheSerPheGlyValValLeuLeuGluLeuThrGlyArg 472
Db 243 ACAGAAAAATCCGATGATATTTCTTTTCGGAGTCATGCTTCTTGAGCTTATTACTGGGCGG 302
Qy 473 ArgProValAspAlaAsnValTyrValAspAspSerLeuValAspTyrAlaArgPro 492
Db 303 CGACCAATTGACACACACCAACCATATATGATGATGACAGCTTGGTTGACTGGCGCAAGGCCA 362

```

```

Qy 493 LeuLeuAsnArgAlaSerGluGlnGlyAspPheGluLeuAlaAspAlaLysMetAsn 512
Db 363 TTACTGATGAGAGCGCTCGAGGATGGTGAATATGATGCTTTGGTGGATCTCTCGGCTCGGA 422
Qy 513 AsnGlyTyrAspArgGluGluMetAlaArgMetValAlaCysAlaAlaAlaCysValArg 532
Db 423 AAGACTTCAATCTTAACGAGATGGCAAGATGATGATGCTGTGCGGTGCGGTGTACGC 482
Qy 533 HisSerAlaArgArgProArgMetSerGlnIleValArgAlaLeuGluGlyAsnVal 552
Db 483 CATTCGACGCTGCTGCGCCACGAATGAGTCAGTCTGCGGTCTCTGGAAGCGCAGCTG 542
Qy 553 SerLeuSerAspLeuAsnGluGlyMetArgProGlyGlnSerAsnValTyrSerTyr 572
Db 543 TCTTTGGAGGACCTTAATGAAGGTGTTGCACTCGCCCATAGCCCTCTTTGGGTCTATAT 602
Qy 573 GlyGlySerThrAspTyrAspSerSerGlnTyrAsnGluAspMetLysLysPheArgLys 592
Db 603 ---AGCAGCTCTGATTACGATTCGGGCGAGTACCAACGAGGACATGCAAGATTTCAGGAG 659
Qy 593 MetAlaLeuGlyThrGlnGluTyrAsnAlaThrGlyGluTyrSerAsnProThrSerAsp 612
Db 660 ACGGCGTTCAATAACAACACCTACACCGAGCAGC---CAATACAGCGCGCCCAACCACTGAA 716
Qy 613 TyrGlyLeuTyrProSerGlySerSerGluGlyGlnThrThrArgGluMet---Glu 631
Db 717 TACGGCCAGGTGCGCTGCTGATCAAGCAGCGAGGCGCCACGAGGAGATGGAGAG 776
Qy 632 MetGlyLysIleLysArgThrGly-----GlnGlyTyrSerGlyProSer 646
Db 777 TCGGGTCGATGAAGAAGAGGTGGCTACAGCGGCTACAGCTCCGGATACAGCGGAGCCTCG 836

RESULT 13
BM358715      666 bp      mRNA      linear      EST 09-JAN-2002
LOCUS         GA_Ea0012D16r Gossypium arboreum 7-10 dpa fiber library Gossypium
DEFINITION    arboreum cDNA clone GA_Ea0012D16r, mRNA sequence.
ACCESSION     BM358715
VERSION       1
KEYWORDS      EST.
SOURCE        Gossypium arboreum
ORGANISM      Gossypium arboreum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
REFERENCE     1 (bases 1 to 666)
AUTHORS       Wing,R.A., Frisch,D., Yu,Y., Main,D., Rambo,T., Simmons,J.,
                Henry,D., Wood,T.C., Leslie,A. and Wilkins,T.A.
TITLE         An integrated analysis of the genetics, development, and evolution
                of the cotton fiber
JOURNAL        Unpublished (2000)
COMMENT       Contact: Wing RA
                Clemson University Genomics Institute
                100 Jordan Hall, Clemson, SC 29634, USA
                Tel: 864 656 7288
                Fax: 864 656 4293
                Email: rwing@clemson.edu
                Total High Quality bases = 646
                Seq primer: TAATACGACTCACTATAGG
                High quality sequence stop: 666.
FEATURES       Location/Qualifiers
                1..666
                /organism="Gossypium arboreum"
                /mol_type="mRNA"
                /strain="AKA"
                /cultivar="8400"
                /db_xref="taxon:29729"
                /clone="GA_Ea0012D16r"
                /tissue_type="Fibers isolated from bolls harvested 7-10
                dpa"
                /lab_host="E. coli"

```

```

/clone_lib="Gossypium arboreum 7-10 dpa fiber library"
/notes="Vector: pBK-CMV; Site_1: EcoRI; Site_2: XhoI"

ORIGIN
Alignment Scores:
Pred. NO.: 7.55e-24 Length: 666
Score: 1018.00 Matches: 191
Percent Similarity: 95.41% Conservat: 17
Best Local Similarity: 87.61% Mismatches: 10
Query Match: 29.48% Indels: 0
DB: 12 Gaps: 0

US-10-086-464-2 (1-647) x BM358715 (1-666)

Qy 247 LeuProProSerProGlyLeuValLeuGlyPheSerLysSerThrPheThrTyrGlu 266
Db 11 CTTCCGGCTCGTCACCTGGTATTTCTTAGTTTCTGAAACACATTTTAGTATGAA 70
Qy 267 GluLeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyPhe 286
Db 71 GAATTAGCGAGACACGATGGCTCTCGGAAGTTAACTTCTTGGACAAGGTGGTTT 130
Qy 287 GlyTyrValHisLysGlyValLeuProSerGlyLysGluValAlaValLysGlnLeuLys 306
Db 131 GGGTACGTACACAAGAGTTCTCCTAATGGGAAGAGTAGCAGTAAAGCACTCAAG 190
Qy 307 ValGlySerGlyGlnGlyGluArgGluPheGlnAlaGluValGluLeuLeuSerArgVal 326
Db 191 GCTGGAAGTGGCAAGCGAGAGAAATTTGAGCTGAAGTTGAGATTCATTAGCCGGTC 250
Qy 327 HisHisArgHisLeuValSerLeuValGlyTyrCysIleAlaGlyAlaLysArgLeuLeu 346
Db 251 CATCACAAACATCTCGTCTCATTCGTCGATCTGATTTCTGGACAAATAAGTCTT 310
Qy 347 ValTyrGluPheValProAsnAsnLeuGluLeuHisLeuHisGlyGluGlyArgPro 366
Db 311 GTTTATGATTTGTTTCCAAACACACCTTGGAGTTTCACTTGCATGGGAAGGGCGACTG 370
Qy 367 ThrMetGluTyrSerThrArgLeuValLeuAlaLeuGlySerAlaLysGlyLeuSerTyr 386
Db 371 ACCATGATTTGCGCAGCAAGGATGAAATTCCTTAGATTCGCAAAAGGACTGGCATAT 430
Qy 387 LeuHisGluAspCysAsnProLysIleLeuHisArgAspIleLysAlaSerAsnIleLeu 405
Db 431 CTTTCATGAGATTTCTCATCTCAAGATCATTCACCGTGATTAAGCGCGCTAATATCTG 490
Qy 407 IleAspPheLysPheGluAlaLysValAlaAspPheGlyLeuAlaLysIleAlaSerAsp 426
Db 491 TTGATTTCAAGTTTGAAGCAAGGTTGCTGATTTTGGACTAGCGAAATTTGCTTCCGAT 550
Qy 427 ThrAsnThrHisValSerThrArgValMetGlyThrPheGlyTyrLeuAlaProGluTyr 446
Db 551 GTCAACACGACGCTCCACAGGATGATGGTACTTTTCGGGTATTTAGCCCTTGAGTAT 610
Qy 447 AlaAlaSerGlyLysLeuThrGluLysSerAspValPheSerPheGlyValVal 464
Db 611 GCTTCAAGTGGAGACTCACTGATAAATCAGATGTTTTCCTTCCTCGGGTCATG 664

RESULT 14
CA072174
LOCUS
DEFINITION SCCAM1004A11.9 AM1 Saccharum officinarum cDNA clone SCCAM1004A11
5', mRNA sequence.
ACCESSION CA072174
VERSION CA072174.1 GI:34924325
KEYWORDS EST.
SOURCE Saccharum officinarum
ORGANISM Saccharum officinarum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
clade; Panicoideae; Andropogoneae; Saccharum.
REFERENCE 1 (bases 1 to 695)
AUTHORS Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.

```

```

TITLE The libraries that made SUCEST
JOURNAL Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
COMMENT Contact: Arruda P
Centro de Biologia Molecular e Engenharia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: parruda@unicamp.br
Clone distribution: clone distribution information can be found
through the Brazilian Clone Collection Center (BCCC) at
http://www.bcccenter.fcav.unesp.br
Plate: 004 row: A column: 11
Seq primer: T7 Promoter Primer.
FEATURES
source
1..695
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCCAM1004A11"
/lab_host="DH10B"
/clone_lib="AM1"
/notes="Organ: Apical meristem and tissues surrounding of
mature plants; Vector: pSport1; Site 1: SalI; Site 2:
NotI; An unidirectional cDNA library generated from
[Apical meristem and tissues surrounding of mature
plants]. cDNA was prepared from polyA+ mRNA using
SuperScript Plasmid System Kit (Invitrogen). The
double-strand cDNAs were fractionated in a separose
CU-2B 40cm-columns and fragments sizing between 0.8 and
1.5 Kb were directionally cloned into the vector. Details
of each source of RNA and library construction can be
obtained at http://sucest.lad.ic.unicamp.br/public"
ORIGIN
Alignment Scores:
Pred. No.: 1.13e-23 Length: 695
Score: 1013.00 Matches: 195
Percent Similarity: 91.85% Conservat: 19
Best Local Similarity: 83.69% Mismatches: 16
Query Match: 29.34% Indels: 3
DB: 13 Gaps: 1
US-10-086-464-2 (1-647) x CA072174 (1-695)
Qy 271 AlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyPheGlyTyrValHis 290
Db 4 GCAACCAATGGGTTTCTGACGCTAATCTGCTGGGCAAGCGGTTTGGGTTTGTTCAC 63
Qy 291 LysGlyValLeuProSerGlyLysGluValAlaValLysGlnLeuLysValGlySerGly 310
Db 64 AAAGGAGTGTACCAATGGCAGAGGTTGCTTTAAGCAGTTAAGAGATGGAAGTGC 123
Qy 311 GlnGlyClnArgGluPheGlnAlaGluValGluLeuLeuSerArgValHisArgHis 330
Db 124 CAGGAGAACCGAGTTTCAGGCAGAGGTTGAGATTATCAGCAGAGTACATCAAGCAT 183
Qy 331 LeuValSerLeuValGlyTyrCysIleAlaGlyAlaLysArgLeuLeuValTyrGluPhe 350
Db 184 CTTGTATCTCTGTGCGCTATTGCTATTTCTGGAGCCAAACAGGTGCTGCTACGAGTTT 243
Qy 351 ValProAsnAsnAsnLeuGluLeuHisLeuHisGlyGlnGlyArgProThrMetGluTyr 370
Db 244 GTTCCAAACATACATACATGGATTCACATTCACATGAAAGAGCAGACCAACCTTGGATGG 303
Qy 371 SerThrArgLeuLysIleAlaLeuGlySerAlaLysGlyLeuSerTyrLeuHisGluAsp 390
Db 304 CCACAAAGACTAAATATGCTCTGGGTTCTGCGAAGGATTGCGATATCTTCATGAAGAT 363
Qy 391 CysAsnProLysIleLeuHisArgAspIleLysAlaSerAsnIleLeuLeuAspPheLys 410
Db 364 TGCCATCTCAAGATCATTCATCGTCGATCAATAAGCCCTCAAAATATTTCTTCTGATCTAAGA 423

```

```

Qy 411 PheGluAlaLysValAlaAspPheGlyLeuAlaLysIleAlaSerAspThrAsnThrHis 430
Db 424 TTTGAAGCTAAGTGGCAGATTTTGGACTTCGAAAATTCACCTTCGTATACAAACACCCAT 483
Qy 431 ValSerThrArgValMetGlyThrPheGlyTyrLeuAlaProGluTyrAlaAlaSerGly 450
Db 484 GTTTCACACAGAGTAATGGGCACATTTGGGTACCTAGCACCTGAGTATGCTGCTTCGGC 543
Qy 451 LysLeuThrGluLysSerAspValPheSerPheGlyValValLeuLeuGluLeuIleThr 470
Db 544 AAGCTCACTGAGAAATCAGATGCTCTCTTTTGGAGTAATGCTCTTTGAGCTAATAACT 603
Qy 471 GlyArgArgProValAspAlaAsnAsnValTyrValAspAspSerLeuValAspTyrAla 490
Db 604 GGGCGGCGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 657
Qy 491 Arg-ProLeuLeuAsnArgAlaSerGluGlnGlyAsp 502
Db 658 AGGGCTTTGATGAATAAAGCAATTTGAGGATGTTAAT 694

RESULT 15
LOCUS BU099573
DEFINITION WHE3309 A09 A172S Chinese Spring wheat drought stressed root cDNA
library_Triticum aestivum cDNA clone WHE3309_A09_A17, mRNA
sequence.
ACCESSION BU099573
VERSION BU099573.1 GI:22547372
KEYWORDS EST.
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooidae; Triticeae; Triticum.
REFERENCE 1 (bases 1 to 715)
AUTHORS Anderson,O.D., Chao,S., Close,T.J., Crossman,C., Fenton,R.D.,
Lazo,G.R., Nguyen H.T., Pham,J., Rausch,C.J., Turuspekov,Y.,
Wilson,C., Woo,J. and Zhang,D.
The structure and function of the expressed portion of the wheat
genomes - Chinese Spring drought stressed root cDNA library
Unpublished (2002)
Contact: Olin Anderson
US Department of Agriculture, Agriculture Research Service, Pacific
West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 5105595773
Fax: 5105595818
Email: oanderson@w.usda.gov
Sequences have been trimmed to remove vector sequence and low
quality sequence with phred score less than 20
Seq primer: SK primer.
FEATURES
Location/Qualifiers
1..715
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/clone="WHE3309_A09_A17"
/tissue_type="root"
/dev_stage="Full tillering stage"
/lab_host="E. coli SOLR"
/clone_lib="Chinese Spring wheat drought stressed root
cDNA library"
/note="Vector: Lambda Uni-ZAP XR, excised phagemid;
Site 1: EcoRI; Site 2: XhoI; Plants were grown under
normal conditions, then drought stressed to 80%, 70% and
60% RWC at Texas Tech University (D. Zhang in Ht Nguyen
lab). Total RNA was prepared separately for roots
collected at the three different drought conditions. Equal
amounts of total RNA were pooled from all three samples,
poly(A) RNA were purified, one cDNA library was made, and
the cDNA clones were in vivo excised to give pBluescript
SK(-) phagemids in the TU Close lab at the University of

```

```

ORIGIN
Alignment Scores:
Pred. No.: 1.04e-22 Length: 715
Score: 983.00 Matches: 184
Percent Similarity: 89.70% Conservative: 25
Best Local Similarity: 78.97% Mismatches: 24
Query Match: 28.47% Indels: 0
Gaps: 0
DB: 13

US-10-086-464-2 (1-647) x BU099573 (1-715)
Qy 276 SerGluAlaAsnLeuLeuGlyGlnGlyPheGlyTyrValHisLysGlyValLeuPro 295
Db 11 TCCGATGCTAATCTTCTCGGGCAAGGTGGCTTTGGATTGTTTCAAAAGGAGTCTGCCA 70
Qy 296 SerGlyLysGluValAlaValLysGlnLeuLysValGlySerGlyGlnGlyGluArgGlu 315
Db 71 GATGGCAGAGAAGTTGCTGTGAAGCAATTAAGAGATGGAAGTGGGCGAGAGCGGTGAG 130
Qy 316 PheGlnAlaGluValGluIleIleSerArgValHisHisArgHisLeuValSerLeuVal 335
Db 131 TTCAGGCAGAGGTTGAGATTATCAGCCGAGTACATCAATAACATCTCGTGACATTTGGTT 190
Qy 336 GlyTyrCysIleAlaGlyAlaLysArgLeuLeuValTyrGluPheValProAsnAsnAsn 355
Db 191 GGTATTATGATTTCTGAAGACAAGAGGTGCTGCTCTATGAGTTTGTTCCTCAATAACACG 250
Qy 356 LeuGluLeuHisLeuHisGlyGluGlyArgProThrMetGluTrpSerThrArgLeuLys 375
Db 251 TTAGAATTCCATATACATGGAAGCGGTGGCAACTATGACTGGCTTCAAGACTACGT 310
Qy 376 IleAlaLeuGlySerAlaLysGlyLeuSerTyrLeuHisGluAspCysAsnProLysIle 395
Db 311 ATTGCTTTGGTCTCGAAGGAGTGGCGTATCTTCACGAAGACTGCCATCCAAAGATC 370
Qy 396 IleHisArgAspIleLysAlaSerAsnIleLeuIleAspPheLysPheGluAlaLysVal 415
Db 371 ATTCACTGTCACATAAAGGCATCAATATTTCTCTGGATTACAGATGTGAAGCTAAGGTG 430
Qy 416 AlaAspPheGlyLeuAlaLysIleAlaSerAspThrAsnThrHisValSerThrArgVal 435
Db 431 GCAGATTTTGGACTTGCAAAAGTTAAACCTCTCGATAATAACACTCATGTGTTCCACCAGAGTA 490
Qy 436 MetGlyThrPheGlyTyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLys 455
Db 491 ATGGGCACATTTGGGTACCTTGCCACAGATATGCTTCTCTGGCAAGCTTAACCTGGCGCGT 550
Qy 456 SerAspValPheSerPheGlyValValLeuLeuGluLeuIleThrGlyArgArgProVal 475
Db 551 TCAGATGCTCTTTCTTTTGGAGTAATGCTTCTCGAGTTAATAACTGGGCGCGTCTCTGTA 610
Qy 476 AspAlaAsnValTyrValAspAspSerLeuValAspTyrAlaArgProLeuLeuAsn 495
Db 611 AGTTCAAAACAAGCGCATATGGATGACAGCTTGGTTGACTGGGCAAGCTTTTGTATGACA 670
Qy 496 ArgAlaSerGluGlnGlyAspPheGlyGlyLeuAlaAsp 508
Db 671 CAAGCACTCGAGGATGGTATATCAGATGCTTTAGTGGAT 709

```

Search completed: April 25, 2004, 00:59:49
Job time : 3628 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 24, 2004, 20:30:41 ; Search time 133 Seconds
(without alignments)
2699.648 Million cell updates/sec

Title: US-10-086-464-2
Perfect score: 3453
Sequence: 1 MSSAPSGTGTGSPSPNST.....REMGKIKRTGGYSGPSL 647

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10086464/runat_23042004_083016_5642/app_query.fasta_1.839
-DB=Issued Patents NA -QMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS-human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORF=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10086464@cgn1_1_85@runat_23042004_083016_5642 -NCFU=6 -ICFU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA.*

- 1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
- 2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
- 3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
- 4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
- 5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq.*
- 6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	756	21.9	3239	4	US-09-228-986-9
2	744.5	21.6	1593	4	US-09-602-472A-1
3	740	21.4	2868	4	US-09-228-986-4
4	712	20.6	1488	4	US-09-579-182-4
5	703.5	20.4	2513	4	US-09-228-986-13
6	678	19.6	2432	4	US-09-228-986-7
7	667.5	19.3	1559	4	US-09-602-472A-3
8	604.5	17.5	4104	3	US-08-881-706-1
9	602	17.4	1224	4	US-09-579-182-3
10	600.5	17.4	3097	4	US-09-228-986-11
11	581.5	16.8	3590	1	US-08-587-889-1
12	581.5	16.8	3590	4	US-09-016-434-1093

13	581.5	16.8	3590	5	PCT-US96-09193-1	Sequence 1, Appli
14	578	16.7	2114	4	US-09-602-472A-5	Sequence 5, Appli
15	575	16.7	2749	1	US-08-265-628-1	Sequence 1, Appli
16	572	16.6	2571	1	US-07-717-331F-9	Sequence 9, Appli
17	572	16.6	2833	1	US-07-717-331F-1	Sequence 1, Appli
18	567	16.4	1554	2	US-08-587-680A-24	Sequence 24, Appli
19	566.5	16.4	2389	4	US-09-228-986-1	Sequence 1, Appli
20	562	16.3	2336	4	US-09-228-986-10	Sequence 10, Appli
21	559.5	16.2	2638	4	US-09-228-986-8	Sequence 8, Appli
22	546	15.8	2749	1	US-07-717-331F-4	Sequence 4, Appli
23	545	15.8	966	1	US-08-447-185-2	Sequence 2, Appli
24	543	15.8	2443	1	US-08-447-185-3	Sequence 3, Appli
25	543	15.7	2686	4	US-09-228-986-3	Sequence 3, Appli
26	531.5	15.4	5733	2	US-08-473-553A-1	Sequence 1, Appli
27	486	14.1	2943	4	US-09-503-922-2	Sequence 2, Appli
28	460.5	13.3	2649	4	US-09-228-986-12	Sequence 12, Appli
29	430	12.5	2568	4	US-09-228-986-2	Sequence 2, Appli
30	401	11.6	1926	4	US-09-249-585A-2	Sequence 2, Appli
31	401	11.6	1926	4	US-09-410-399-3	Sequence 3, Appli
32	401	11.6	2580	3	US-09-050-863-2	Sequence 2, Appli
33	401	11.6	2580	4	US-09-359-081-2	Sequence 2, Appli
34	401	11.6	5452	2	US-09-130-114-1	Sequence 1, Appli
35	401	11.6	8705	4	US-09-647-344A-14	Sequence 14, Appli
36	401	11.6	9600	3	US-08-910-647-1	Sequence 1, Appli
37	401	11.6	9600	4	US-09-620-925-1	Sequence 1, Appli
38	401	11.6	10596	1	US-07-884-811-15	Sequence 15, Appli
39	401	11.6	10596	1	US-07-885-971-15	Sequence 15, Appli
40	401	11.6	10596	1	US-08-087-783A-15	Sequence 15, Appli
41	401	11.6	10596	1	US-08-194-088B-15	Sequence 15, Appli
42	401	11.6	10596	2	US-08-194-087-15	Sequence 15, Appli
43	401	11.6	10596	5	PCT-US93-04648-15	Sequence 15, Appli
44	401	11.6	16080	4	US-09-724-566A-48	Sequence 48, Appli
45	395.5	11.5	4403765	3	US-09-103-840A-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1

US-09-228-986-9
; Sequence 9, Application US/09228986
; Patent No. 6359198
; GENERAL INFORMATION:
; APPLICANT: Nieuwenhuizen, Timothy
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signalling
; FILE REFERENCE: 11000/1020
; CURRENT APPLICATION NUMBER: US/09/228,986
; CURRENT FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 130
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9
; LENGTH: 3239
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-09-228-986-9

Alignment Scores: 1.88e-23 3239
Pred. No.: 756.00 Matches: 190
Score: 51.93% Conservative: 92
Percent Similarity: 34.99% Mismatches: 132
Best Local Similarity: 21.89% Indels: 130
Query Match: 4 Gaps: 17
DB:

US-10-086-464-2 (1-647) x US-09-228-986-9 (1-3239)

Qy	89	ProProalaProVal-ThrProThrArgAsnProProSerValProGlyProPr	108
Db	1759	CTCCCAACCAATTCGCCCTATTACTTCATCCCTCCCTATGCTTTT	1811
Qy	108	oserAsnProSerArgGluGlySerProArgProProSerProProSe	128

```
Db 1812 AGACAAT-----GGAGGAACC-----1827
Qy 128 rProSerSerAspGlyLeuSerThrGlyValValGlyLeuAlaIleGlyValAl 148
Db 1828 -----GCCATAAGCAAAAGGTGTGATAGTTGGATCGTATTGGCGCAGGT 1874
Qy 148 aLeuLeuValIleValThrLeuLeuCysLeuLeuGlyValSerLeuValAla 168
Db 1875 TCTGGTCTTCTGGCCTTGTGTATTAGGGTTATATGCTATTCGACAAAAGAACGGGGGA 1934
Qy 168 uGluAspAlaTyrTrpValProPro-----ProProGlyProLysAl 184
Db 1935 GAAA---GCTCTCAGTTGACACACCTTCGCATCTCGGCACCCAGTGGGAAGATAG 1991
Qy 184 aGlyGlyProTyrGlyGlyGlnGlnIntrpArgGlnGlnAsnAlaThrProPse 204
Db 1992 CGGAGGAGCGCCACACACTGAAAGGACGACGATGG-----2025
Qy 204 rAspHisValValThrSerLeuProProProLysAlaProSerProProArgGlnPr 224
Db 2025 -----2025
Qy 224 oProProProProProPheMetSerSerSerGlyGlySerAspTyrSerAspAr 244
Db 2025 -----2025
Qy 244 gProValLeuProProProSerProGlyLeuValLeuGlyPheSerLysSerThrPheTh 264
Db 2026 -----TTCTC 2030
Qy 264 rTyrGluGluLeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGl 284
Db 2031 CTATGATGAACCTTAAGAGGTGCACCAATAATTTCTCCGATAGCAATGAATTAGGCTTCGG 2090
Qy 284 yGlyPheGlyTyrValHisLysGlyValLeuProSerGlyLysGluValAlaValLysGl 304
Db 2091 AGGATACGGAAGGTGTACAGGGAGTCTTCTGTAGTGTCTATATATTAGCAATCAAAAG 2150
Qy 304 nLeuLysValGlySerGlyGlnGlyGluArgGluPheGlnAlaGluValGluIleLese 324
Db 2151 AGCTCAGCAGGGTTCGATCAGAGGTGCAACCGAGTTCAAGACAGAAATCGAGCTGCTTC 2210
Qy 324 rArgValHisHisArgHisLeuValSerLeuValGlyTyrCysIleAlaGlyAlaLysAr 344
Db 2211 CGCGGTTTCATCACAAGAACTCTGTGTGGCTCATAGGATCTGTTCGAGCAAGAGAGCA 2270
Qy 344 gLeuLeuValTyrGluPheValProAsnAsnAsnLeuLeuHisGlyGlyGluGl 364
Db 2271 GATGTTGTTCTATGATATATATGCTTAACGGACGCTCAGGGATAGCTTGACAGAAATC 2330
Qy 364 yArgProThrMetGluTrpSerThrArgLeuLysIleAlaLeuGlySerAlaLysGlyLe 384
Db 2331 AGGCATTATCTTGATGGAAGAGAGGCTTCGTATAGCTCTAGGTTCCGCTAGAGACT 2390
Qy 384 uSerTyrLeuHisGluAspCysAsnProLysIleIleHisArgAspLysAlaSerAs 404
Db 2391 AGCTTATCTGCAAACTCGCGAATCTCTCCAAATTATCCACAGAGATGTCAAAGTCCACCAA 2450
Qy 404 nIleLeuLeuAspPheLysPheGluAlaLysValAlaAspPheGlyLeuAlaLysIleAl 424
Db 2451 TATCTTGTGAGCAACATCTGACGGCCAAAGTCCGGATTTCCGTTTGTCCAACTGGT 2510
Qy 424 aSerAspThrAsnThr---HisValSerThrArgValMetGlyThrPheGlyTyrLeuAl 443
Db 2511 ATCGGACAGCGGAAGGGGACGTTTCGACCAAGTGAAAGGCACGCTGGGCTATTGGA 2570
Qy 443 aProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerSerAspValPheSerPheGlyVa 463
Db 2571 TCCCGAATACTACATGATAGTCAACAGCTGACAGAAAGAGCGATGTGTACAGCTTCGGGGT 2630
Qy 463 lValLeuLeuGluLeuIleThrGlyArgArgProValAspAlaAsnAsnValTyrValAs 483
Db 2631 GGTATGCTTGAGTCTCATCTGCAAGCAACCGATTGAGAGGCAAGTATGTCGTCGG 2690
```

```
Qy 483 pAspSerLeuValAspTrpAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPh 503
Db 2691 CGAG-----ATTCGACCGCCATCGACAAGAACACCCAG-----GACTA 2729
Qy 503 eGluGly-----LeuAlaAspAlaLysMetAsnAsn---GlyTyrAspArgGluGl 519
Db 2730 CTACGGCGTGAAGAAATGATGGACCCGTCATGAGGAGCATGGGCTAC---CTCGTCGG 2786
Qy 519 uMetAlaArgMetValAlaCysAlaAlaAlaCysValArgHisSerAlaArgArgArgPr 539
Db 2787 GTTCAGCAGGTTCTTGATTTGGCGATGCGATGTCGAGGAGTCGCTCGGACCGCCC 2846
Qy 539 oArgMetSerGlnIleValArgAlaLeuGluGlyAsnValSerLeuSerAspLeuAsnGl 559
Db 2847 CACAATGAGCGAGTGTGAGGCGATCGAGCATGTTGTCAG-----AACGA 2894
Qy 559 uGlyMetArgProGlyGlnSerAsnValTyrSerSerTyrGlyGlySerThrAspTyrAs 579
Db 2895 TGGGATACACACCAACTCGACGTCGCGATCGCTCGTCG-----GCGACGGACTTGG 2945
Qy 579 pSerSerGln-----TyrAsnGluAspMetLysLysPheArgLysMe 593
Db 2946 GTCACGAGGGCGCTCTCGGCATCGGTACACAGATCCCTTACCAG-----2994
Qy 593 tAlaLeuGlyThrGlnGluTyrAsnAlaThrGlyGluTyrSerAsnProThrSerAspTy 613
Db 2995 -----AAGGAAGTTAGCTATAGCATTCCTTTGATTATAGT-----GGTGGATA 3038
Qy 613 rGlyLeu 615
Db 3039 TGGACTA 3045

RESULT 2
US-09-602-472A-1
; Sequence 1, Application US/09602472A
; Patent No. 6608240
; GENERAL INFORMATION:
; APPLICANT: Bidney, Dennis L.
; APPLICANT: Hu, Xu
; APPLICANT: Li, Guihua
; TITLE OF INVENTION: Sunflower Disease Resistance Genes
; FILE REFERENCE: 35718/200630
; CURRENT APPLICATION NUMBER: US/09/602,472A
; CURRENT FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/140,876
; PRIOR FILING DATE: 1999-06-24
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1593
; TYPE: DNA
; ORGANISM: Helianthus annuus
; FEATURE:
US-09-602-472A-1

Alignment Scores:
Pred. No.: 2,9e-23 Length: 1593
Score: 744.50 Matches: 172
Percent Similarity: 57.07% Conservative: 66
Best Local Similarity: 41.25% Mismatches: 116
Query Match: 21.56% Indels: 63
DB: 4 Gaps: 9

US-10-086-464-2 (1-647) x US-09-602-472A-1 (1-1593)
Qy 235 SerSerGlyGlySerAspTyrSerAspArgProValLeuProProSerProGlyLeu 254
Db 181 TCGCGGGTTCGAAAGACGGGTCGACGGTCAGATT-----216
Qy 255 ValLeuGlyPheSerLysSerThrPheThrTyrGluGluLeuAlaArgAlaThrAsnGly 274
Db 217 -----GCTGCTCATACATTCCATTTCGGGAGCTTGCAGCTGCAACAAACAA 264
```



```
QY 287 GlyTyrValHisLysGlyValLeuProSerGlyLysGluValAlaValLysGlnLeuLys 306
Db 1188 GGAAGGTTGTAACAAAGCGCGCTTCGAGATGTTCTTGGTGGCTGTAAACCGTCTGAAG 1247
QY 307 ---ValGlySerGlyGlnGlyGluArgGluPheGlnAlaGluValGluLeuLeuSerArg 325
Db 1248 GAAGAGCTACACCGGGTGGAGAGTTGCAGTTTCAAACAGAGTGGAGATGATTAAGCATG 1307
QY 326 ValHisHisArgHisLeuValSerLeuValGlyTyrCysIleAlaGlyAlaLysArgLeu 345
Db 1308 GCAGTACATAGGAACCTCTTCGACTACGTGATTCGATGACACACCCACTGAACGGCTG 1367
QY 346 LeuValTyrGluPheValProAsnAsnLeuGlu-----LeuHisLeuHisGlyGlu 363
Db 1368 CTTGTTTATCCTACATGCGCAAGTGAAGTGTGCTTCATCCCTACAGAGAGGCGCA 1427
QY 364 GlyArgProThrMetGluTyrSerThrArgLeuLysIleAlaLeuGlySerAlaLysGly 383
Db 1428 AATGACCACCTTAGATGGCCAACTCGAAGCGCATAGATTGGGTCTGCAAGAGGT 1487
QY 384 LeuSerTyrLeuHisGlyAspCysAsnProLysIleLeuHisArgAspIleLysAlaSer 403
Db 1488 CTCTCCTACTTGCATGATCATTTGTGATCTTAAGATTATTCACCGGATGTCAAGCGTCT 1547
QY 404 AsnIleLeuIleAspPheLysPheGluAlaLysValAlaAspPheGlyLeuAlaLysIle 423
Db 1548 AACATCTTACTTGGATGAAGATATAGAGCAGTGTGGGGGATTTTGGCTTGGCAAACTT 1607
QY 424 AlaSerAspThrAsnThrHisValSerThrArgValMetGlyThrPheGlyTyrLeuAla 443
Db 1608 ATGGATTATAGGACACACATGTTACGACGCTGTTCTGTGAACCATTTGGCCACATAGCA 1667
QY 444 ProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPheSerPheGlyVal 463
Db 1668 CTTGAGTACCTTCTACTGGAAGTCTTCGGAAGACAGACGATATTGGATATGGAATC 1727
QY 464 ValLeuLeuGluLeuIleThrGlyArgArgProValAlaAspAlaAsnValTyrValAsp 483
Db 1728 ATGTTGCTGGAACTTATTACCGGCAACCGGCATTTTGACCTTGCACGTTTAGCAAAATGAT 1787
QY 484 AspSer-----LeuValAspTyrAlaArgProLeuLeuAsnArgAlaSerGluGlnGly 501
Db 1788 GATGATGTCATGTTGCTTACTGCTGATGGTTAAAGCCCTACTA-----AAAGAGAGA 1835
QY 502 AspPheGlyGlyLeuAlaAspAlaLysMetAsnAsnGlyTyrAspArgGluGluMetAla 521
Db 1836 AGGCTTGATATGCTAGTTGATCCTGATCTTAAGAACAAATTATGTTGAAGCAGAGGTGGA 1895
QY 522 ArgMetValAlaCysAlaAlaCysValArgHisSerAlaArgArgArgProArgMet 541
Db 1896 CAACATTATCAAGTTGCAATTACTTTGTACACAAGGGTCAACCAATGGATAGACCAAGATG 1955
QY 542 SerGlnIleValArgAlaLeuGluGlyAsn-ValSerLeuSerAspLeuAsnGluGlyMet 561
Db 1956 TCTGAGTGTGAGGATGTTGGAAGGGATGGCTTAGCTGAGAGATGGAGGAATGGCAA 2015
QY 561 tArgProGlyGlnSerAsnValTyrSerSerTyrGly 573
Db 2016 AAGG-----TGGAAGTCTGTACGGA 2034
```

RESULT 4

```
US-09-579-182-4
; Sequence 4, Application US/09579182
; Patent No. 6500628
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN KINASE AND
; CURRENT APPLICATION NUMBER: US/09/579,182
; CURRENT FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 7
```

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 4

; LENGTH: 1488

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-579-182-4

Alignment Scores:

Pred. No.:	6,04e-22	Length:	1488
Score:	712.00	Matches:	178
Percent Similarity:	53.07%	Conservative:	73
Best Local Similarity:	37.63%	Mismatches:	160
Query Match:	20.62%	Indels:	62
DB:	4	Gaps:	14

US-10-086-464-2 (1-647) x US-09-579-182-4 (1-1488)

```
QY 205 AspHisValValThrSerLeuProPro-----ProProLysAlaProSer 219
Db 22 GAACACAGATAAATCCCTTCGCTCCGCTATATTCTTAATCCGATAATCAAGTCT 81
QY 220 ProProArgGlnProProProProProProProProPheMetSerSerGlyGlySer 239
Db 82 -----CGACGAGGTCTCT----- 93
QY 240 AspTyrSerAspArgProValLeuProProProSerProGlyLeuValLeuGlyPheSer 259
Db 94 -----GAACAGAGAAGGAGCTAACTGCTCCAAAAGAGGGCTACTGCGCATATTGCT 147
QY 260 LysSerThrPheThrTyrGluGluLeuAlaArgAlaThrAsnGlyPheSerGluAlaAsn 279
Db 148 GCACAAACCTTACTTTCCGAGAGCTTAGTCCGCCACTAAACACTTTTCGACCGGAATGT 207
QY 280 LeuLeuGlyGlnGlyPheGlyTyrValHisLysGlyValLeu---ProSerGlyLys 298
Db 208 CTTCTTGAGAGAAGGAGGTTTCGACGCTGTTTACAAAGGTCTCTAGAGACCAACAGGACAG 267
QY 299 GluValAlaValLysGlnLeuLysValGlySerGlyGlnGlyGluArgGluPheGlnAla 318
Db 268 ATAGTAGCTGTAAACAGCTTGTGTAACCGGTCTCAAGAGGAACAGAGAGTTTCTTGTA 327
QY 319 GluValGluIleIleSerArgValHisHisArgHisLeuValSerLeuValGlyTyrCys 338
Db 328 GAGGTTCTTAGCTGAGGCTTCTGCATCATCCCAATCTTGTGAATTTGATTGTTATTGT 387
QY 339 IleAlaGlyAlaLysArgLeuLeuValTyrGluPheValProAsnAsnAsnLeuGluLeu 358
Db 388 GCTGATGGGACCAAGCGCTCTTCTGTGTATGAGTATATGCCACTAGGATCATTCGAGGAT 447
QY 359 HisLeuHisGly-----GluGlyArgProThrMetGluTyrSerThrArgLeuLysIle 376
Db 448 CATCTACACGATCTTCCACCTGTATAAGAGGCTCTAGACTGGAGTACTAGTAATGACAATA 507
QY 377 AlaLeuGlySerAlaLysGlyLeuSerTyrLeuHisGluAspCysAsnProLysIleIle 396
Db 508 GCGCAGAGCAGACAAAGGACTGAGTATCTGATGATAAAGCGAATCCCGCTGTGATC 567
QY 397 HisArgAspIleLysAlaSerAsnIleLeuLeuAspPheLysPheGluAlaLysValAla 416
Db 568 TACAGAGACCTGAAATCATCCAACTTCTCTCGGTGATGGCTATCACCCAAAGTTATCT 627
QY 417 AspPheGlyLeuAlaLysIleAlaSer---AspThrAsnThrHisValSerThrArgVal 435
Db 628 GATTTTGGTTAGCTAAGTTAGTGGTCCCGTGGGCGATATAACACATGTGTCAACTCGTGTG 687
QY 436 MetGlyThrPheGlyTyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLys 455
Db 688 ATGGGCACATATGTTTATTTGTCACCGGAAATATGCCATGACAGGGCAACTCACATTGAA 747
QY 456 SerAspValPheSerPheGlyValValLeuLeuGluLeuIleThrGlyArgProVal 475
Db 748 TCCGATGTTTATAGCTTTGGGGTTGTGTTTCTCGAGCTCATCAGGGGTGCAAAAGCTATT 807
```



```
Qy 476 AspAlaAsnAsnValTyrValAspSerLeuValAspTyrAlaAspProLeuLeuAsn 495
Db 808 GATAATGCTCGAGCACCAGGAGGACACAACTTGTGCGCATGGCTAGGCGGTGTTTC--- 864
Qy 496 ArgAlaSerGluGlnGlyAspGluGlyLeuAlaAspAlaLysMetAsnAsnGlyTyr 515
Db 865 -----AAAGATCGTAGAAGTTTCCGAGATGGCGGATCCATCGCTGCAAGGCGGTAT 918
Qy 516 AspArgGluGluMetAlaAlaCysAlaAlaCysValArgHisSerAla 535
Db 919 CCAATGCGTGTATATCAAGCACTTCAGTTCGAGCAATGTGTTACAGGAACAGCA 978
Qy 536 ArgArgProArgMetSerGlnLeuValArgAlaLeuGluGlyAsnValSerLeuSer 555
Db 979 GCGACAGACCACTGATTGGCGACGTGGTGACAGCTCTTAACATATTAGCTTCGCAACG 1038
Qy 556 ---AspLeuAan-----GluGlyMetArgProGlyGlnSerAsnValTyr--- 569
Db 1039 TTTGACCCAAAGCCACCAAGCGGTCAAAACAGTAGAAGTGGGAGTGGGCCCACTTTATC 1098
Qy 570 -----SerSerTyrGlyGlySerThrAspTyrAsp----- 579
Db 1099 AGAACAGGATGATCGAGAGAGCTTGGAGATGGGAGTAGCTGGATGCTTCGCAGAG 1158
Qy 580 -----SerSerGlnTyrAsnGluAspMetLysLysPheArgLysMet 593
Db 1159 ACTCGAGTGGTTAGGCTCACCAGCCACTCACAAGAACTCTCCTGATTACAGAAGAAG 1218
Qy 594 AlaLeuGlyThrGlnGlyTyrAsnAlaThrGlyGluTyrSerAsnProThrSer--- 611
Db 1219 GATATG---GTGAGGGAAGTCAATGCAGGATCAGAAGGTGGGAGGAGGAGGCGGG 1275
Qy 612 -----AspTyrGlyLeuTyrProSerGlySerSerGluGlyGlnThrArg 628
Db 1276 TCAGGTAGAAATCGGGATTA-----ACCGATTGGAAGGCAAGATCACAG 1323
Qy 629 GluMetGluMetGlyLysLeuLysArgThrGlyGlnGly 641
Db 1324 AGAGGGAGCCGCGGAGTGTGGGAGATCATCGAGAGGC 1362
RESULT 5
US-09-228-986-13
; Sequence 13, Application US/09228986
; Patent No. 6359198
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signalling
; FILE REFERENCE: 11000/1020
; CURRENT APPLICATION NUMBER: US/09/228,986
; CURRENT FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 130
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 13
; LENGTH: 2513
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-228-986-13
Alignment Scores:
Pred. No.: 2,21e-21 Length: 2513
Score: 703.50 Matches: 178
Percent Similarity: 49.52% Conservative: 81
Best Local Similarity: 34.03% Mismatches: 179
Query Match: 20.37% Indels: 85
DB: 4 Gaps: 14
US-10-086-464-2 (1-647) x US-09-228-986-13 (1-2513)
Qy 84 ThrThrProGlySerProAlaProValThrProThrArgAsnProProSer 103
Db 795 ACTGGTCTGCTCCTCCTCCGCCCACTATATCTCCACT----- 833
```

```
Qy 104 ValProGlyProProSerAsnProSerArgGluGlySerProArgProProSerSer 123
Db 834 -----CCACCTTCTTAC-----AATCGGCCACCCCAAGTCC 863
Qy 124 ProSer-----ProProSerProSerSerAspGly-----Leu 134
Db 864 TCAAAATGCTGCTCCCTTCATCAGGTGGCTCAAGGCTGGGAACAGCAATAAGAAATCTCTG 923
Qy 135 SerThrGlyValValValGlyLysAlaLysGlyValAlaLeuLeuValLeuValThr 154
Db 924 AGTGGTGGTCCATAGTGGGTATATAATTTTCAGATTATTTTTCAGTGTGTGCTGCTATA 983
Qy 155 LeuLysCysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 174
Db 984 TTAGGAGTTATTTATATGACGTAACTCTCTAGAGAGAGCAGGAT----- 1031
Qy 175 ProProProProProGlyProLysAlaGlyProTyrGlyGlyGlnGlnGln 194
Db 1032 -----GAAGAAAACTA 1043
Qy 195 TrpArgGlnGlnAsnAlaThrPro-----ProSerAspHisValValThrSerLeuPro 212
Db 1044 AGCAATCGTGTGCTTTTCACCCCTTCATCTCCCTCGATGCTGAATATTATTGAAAGAGAGT 1103
Qy 213 ProProProLysAlaProSerProArgGlnProProProProProProProPhe 232
Db 1104 CCAGAGCAAAAGTCACTCATCACCTCTTGAATCGCTCTTAAGCTCCCTCTTCGAA 1163
Qy 233 MetSerSerSerGly-----GlySerAspTyrSerAspArgProValLeu 247
Db 1164 CGCAACAAGTCTACAGGGGACAAAGGCTTCGGAAGTATTTTTCAGTAGAGAGACTAAA 1223
Qy 248 ProProProSerProGlyLeuValLeuGlyPheSerLysSerThrPheThrGluGlu 267
Db 1224 AACCCA-----ATATCAGCAACTGAATATTCTATTTCAGAC 1259
Qy 268 LeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyPheGly 287
Db 1260 CTGCAATATGCAACAAATAGTTTATGTCAGATATCTTATTCGGGGGGTCTCTTGA 1319
Qy 288 TyrValHisLysGlyValLeuProSerGlyLysGluValAlaValLysGlnLeuLysVal 307
Db 1320 CGAATCTACCGAGCAGAGTTCAGAGTGAAGAGATTTTGGCAGTGAAGAAATTTGACACT 1379
Qy 308 Gly-----SerGlyGlnGlyGluArgGluPheGlnAlaGluLysLeuSerArg 325
Db 1380 TCTACGCTGCTCCCTACAAAGCCCTGAAGACTTCTCGACCGCAGTATCTAATATATCGGC 1439
Qy 326 ValHisHisArgHisLeuValSerLeuValGlyTyrCysLysAlaGlyAlaLysArgLeu 345
Db 1440 CTACATCATCTTAACATTACAGAACTAGTGGTTATTGTCACAGAACATGACATACCTT 1499
Qy 346 LeuValTyrGluPheValProAsnAsnAsnLeu-----GluLeuHisLeuHisGlyGlu 363
Db 1500 CTTGTGTATGAATATTTTCGACAATGGATCACTCTATCAGCGTATTGTCACATGGCAGATGAG 1559
Qy 364 GlyArgProThrMetGluTrpSerThrArgLeuLysLysLysLysLysLysLys 383
Db 1560 ACTACTAGAAATTTGCTTGGAACTTCGTGTAAGAGTTCGGTGGGTTCAGCTCGAGTT 1619
Qy 384 LeuSerTyrLeuHisGluAspCysAsnProLysLysLysLysLysLysLysLysLys 403
Db 1620 TTAGAGTATTGTCATGAGATTTGCTCTCCATCTATTGTCATATAAAAAATTCAGTCTCT 1679
Qy 404 AsnLeuLeuLeuAspPheLysPheGluAlaLysValAlaAspPheGlyLeuAlaLysLeu 423
Db 1680 AATATTTTGTGATGATGATTTCAACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1733
Qy 424 AlaSerAspThrAsnThrHisValSerThrArgValMetGlyThrPheGlyTyrLeuAla 443
Db 1734 GCCCTCAATCAAAATCTGAGCGCTCAGGTTTTCAGGTTTTCAGGTTTTCAGGTTTTCAGGTT 1793
```

```
Qy 444 ProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPheSerPheGlyVal 463
Db 1794 CCTGAATATGTCATGTCAGGAATCTATACAATGAAGACCGATGTGTATAGTTTGGAGTG 1853
Qy 464 ValLeuLeuGluLeuLeuThrGlyArgArgProValAspAlaAsnValTyrValAsp 483
Db 1854 GTATGCTTGAGCTTTTCAGCGCGGAGCCCTTGGATAGTTCAAGAACAGGTCCAGAA 1913
Qy 484 AspSerLeuValAspTTPAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPhe 503
Db 1914 CAGTCATTAGTAAGATGGGCCACCCTCAGCTTCAT-----GATATT 1955
Qy 504 GluGlyLeuAla-----AspAlaLysMetAsnAsnGlyTyrAspArgGluGluMet 520
Db 1956 GATGCATTGGCAAGAGATGGTTCATCCAGCTTTGAAGGGAGGATATCTCGCAAAATCTCTC 2015
Qy 521 AlaArgMetValAlaCysAlaAlaCysValArgHisSerAlaArgArgProArg 540
Db 2016 TCGCGCTTGTGATATTATTGCCCTCTGCATTCAGCTGACACAGAAATTCGTCCTCCA 2075
Qy 541 MetSerGlnIleValArgAlaLeuGluGlyAsnValSerLeuSerAspLeuAsnGluGly 560
Db 2076 ATGCTGAAGTGGTCAAGCATTTGGTTCGTATGATGCAACGGCTAGCTCAATAAGAGG 2135
Qy 561 MetArgProGlyGlnSerAsnValTyrSerSerTyrGlyGlySerThrAspTyrAspSer 580
Db 2136 ATGACAGAGATGAA-----ACTGCAGACCAAGATCCT 2168
Qy 581 SerGlnTyr 583
Db 2169 CGAGATTAT 2177

RESULT 6
US-09-228-986-7
; Sequence 7, Application US/09228986
; Patent No. 6359198
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Niels
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signalling
; FILE REFERENCE: 11000/1020
; CURRENT APPLICATION NUMBER: US/09/228,986
; CURRENT FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 130
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 2432
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-228-986-7

Alignment Scores:
Pred. No.: 2,44e-20 Length: 2432
Score: 678.00 Matches: 147
Percent Similarity: 63.05% Conservative: 68
Best Local Similarity: 43.11% Mismatches: 104
Query Match: 19.64% Indels: 22
DB: 4 Gaps: 8

US-10-086-464-2 (1-647) x US-09-228-986-7 (1-2432)

Qy 254 LeuValLeuGlyPheSerLysSerThrPheThrTyrGluGluLeuAlaArgAlaThrAsn 273
Db 1046 ATTTGCTAGGCGCATTTCAAG---AGATTCTCAATCAAGAGATTACGGTGTTCACATAAT 1102
Qy 274 GlyPheSerGluAlaAsnLeuLeuGlyGlnGlyGlyPheGlyTyrValHisGlyVal 293
Db 1103 AATTTAGTACCAAGATATTATTAGGATGAGGATATGAATGTCTATAAAGATTC 1162
Qy 294 LeuProSerGlyLysGluValAlaValLysGlnLeuLysValGlySer---GlyGlnGly 312
Db 1163 CTACAAGATGGCACTATAGCAATAAAGTTGAAGATGTTAATGTGGAGGAGGA 1222
```

```
Qy 313 GluArgGluPheGlnAlaGluValGluIleLeuSerArgValHisHisArgHisLeuVal 332
Db 1223 GAAATTCAAATTTCAACAGAGAGTGAATGATACAGCTTGGCTGTGCATAGAACCTATTA 1282
Qy 333 SerLeuValGlyTyrCysIleAlaGlyAlaLysArgLeuLeuValTyrGluPheValPro 352
Db 1283 CGATTGTATGATTTTGCACAACTCCAGAGAGGCTTCTGGTCTATCCCTACATGCCA 1342
Qy 353 AsnAsnLeuGlu-----LeuHisLeuHisGlyGluGlyArgProThrMetGluTyrP 370
Db 1343 AATGGAAGTGGCGCTCTTGTCTTAGAGATCATATTAATGAAAAGCTTGCCTCGACTGG 1402
Qy 371 SerThrArgLeuLysIleAlaLeuGlySerAlaLysGlyLeuSerTyrLeuHisGluAsp 390
Db 1403 CCTACTCGCAAGCTATAGCCCTTGGAGCAGCTAGGGAGCTGTATATTTGCATGAGCAA 1462
Qy 391 CysAsnProLysIleIleHisArgAspIleLysAlaSerAsnIleLeuIleAspPheLys 410
Db 1463 TGTGATCCCAGATATTATCCCGGATGTGAACAGCAATAATATATTACTGGATGAATAT 1522
Qy 411 PheGluAlaLysValAlaAspPheGlyLeuAlaLysIleAlaSerAspThrAsnThrHis 430
Db 1523 TTTGAAGCTGTGTGGAGATTTTGGGTAGCAAGCTCTTGGATCACAGGGATTCATCAT 1582
Qy 431 ValSerThrArgValMetGlyThrPheGlyTyrLeuAlaProGluTyrAlaAlaSerGly 450
Db 1583 GTGACTACTGTGTTCGAGGCGGTAGGTACATATGCCCAAGATACCTTTTCAACGGGA 1642
Qy 451 LysLeuThrGluLysSerAspValPheSerPheGlyValValLeuLeuGluLeuIleThr 470
Db 1643 CAATCTTCAGAGAAACTGATGATTTTGGCTTTGGGATATTACITGTTGGAATCATTTACA 1702
Qy 471 GlyArgArgProValAsp-----AlaAsnAsnValTyrValAspAspSerLeu 486
Db 1703 GGACAAAGGCTTTAGATTTTGGCAGGCTGCAAAAGCAAAAAGTTGTA-----ATG 1753
Qy 487 ValAspTrpAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGluGlyLeu 506
Db 1754 CTGATTTGGGTAAAGAGCTT-----CATCAAGAGAGAAAGTTGCACCTCCTT 1801
Qy 507 AlaAspAlaLysMetAsnAsnGlyTyrAspArgGluGluMetAlaArgMetValAlaCys 526
Db 1802 GCTGATAAAGATCTTAAGGGCAATTTGATAGATTGAGCTGAAGAGATGTTTCAGGTT 1861
Qy 527 AlaAlaCysValArgHisSerAlaArgArgProArgMetSerGlnIleValArg 546
Db 1862 TCTTTGTATGCACCAATTTTCAGCTTGGACATCGTCCAAAATGTGTGATGTTTGTAGA 1921
Qy 547 AlaLeuGluGlyAsnValSerLeuSerAspLeuAsnGluGlyMetArgProGlyGlnSer 566
Db 1922 ATGTTGGAAGGTGAT---GGATTGACAGAACGGTGGGAAACATTTGCAAAAAATTGAAACC 1978
Qy 567 AsnValTyrSer-----SerTyrGlyGlySerThrAspTyrAspSer 580
Db 1979 CCCCATACAGAGTAAGTACAGATACCCATATATATCCGAGTTGTTGGAAGAAATTTCT 2038
Qy 581 Ser 581
Db 2039 TCT 2041

RESULT 7
US-09-602-472A-3
; Sequence 3, Application US/09602472A
; Patent No. 6608240
; GENERAL INFORMATION:
; APPLICANT: Bidney, Dennis L.
; APPLICANT: Hu, Xu
; APPLICANT: Li, Guihua
; TITLE OF INVENTION: Sunflower Disease Resistance Genes
; FILE REFERENCE: 35718/200630
; CURRENT APPLICATION NUMBER: US/09/602,472A
; CURRENT FILING DATE: 2000-06-23
```



```
Qy 236 SerGlySerAspTyrSerAspArgProValLeuProProSerPro-----Gly 253
Db 364 AAGACGGATCGAGCGATCTCTGCTAATCCTTTGACAGCTCCATCTCTTCTGCTGGT 423
Qy 254 LeuVal-----LeuGlyPheSerLysSerThrPheThrTyrGluGluLeuAla 269
Db 424 CTTCTCGAGTTTCTCACCTTGGATGGGA---CATTTGGTTCACTCTTAGAGATCTTCAG 480
Qy 270 ArgAlaThrAenGlyPheSerGluAlaAenLeuLeuGlyGlnGlyGlyPheGlyTyrVal 289
Db 481 ATGCTACTAATCAGTTTCAAGGATATATATCATCGGTGATGGTATGGAGTTGTT 540
Qy 290 HisLysGlyValLeuProSerGlyLysGluValAlaValLysGlnLeuLysValGlySer 309
Db 541 TACCGCGTAACTTGTAAATGGTACTCTCTGCTGTTAAAGATTGCTCAACAATTAA 600
Qy 310 GlyGlnGlyGluArgGluPheGlnAlaGluValGluLeuLeuSerArgValHisArg 329
Db 601 GGACAAGCTGATAAAGACTTCAGAGTTGAAGTTGAAGCTATAGGTTCAGTTTCGACATAAA 660
Qy 330 HisLeuValSerLeuValGlyTyrCysIleAlaGlyAlaLysArgLeuLeuValTyrGlu 349
Db 661 AACTTGTGCCCCCTTCGCGATATTGTATGGAAGAACGACGAGG----- 705
Qy 350 PheValProAenAenAenLeuGluLeuHisLeuHisGlyGluGlyArgProThrMetGlu 369
Db 705 ----- 705
Qy 370 TrpSerThrArgLeuLysIleAlaLeuGlySerAlaLysGlyLeuSerTyrLeuHisGlu 389
Db 706 -----CTCGGTACTCTTCAGGAG 723
Qy 390 AspCysAenProLysIleIleHisArgAspIleLysAlaSerAenIleLeuIleAspPhe 409
Db 724 CGGATTGAGCCAAAGTGTGCGACAGACATTAAGTCTAGTAACATTCTGATTGATGAC 783
Qy 410 LysPheGluAlaLysValAlaAspPheGlyLeuAlaLysIleAlaSerAspThrAsnThr 429
Db 784 AAATTCAATTCTAANAATTCGACTTTGGACTTGTCTAAACTACTTGTGTGTGATAAGAGT 843
Qy 430 HisValSerThrArgValMetGlyThrPheGlyTyrLeuAlaProGluTyrAlaAlaSer 449
Db 844 TTTATAACTACTAGAGTTATGGTACCTTCGGTTACGTAGTCCAGAGTATGCCAATTCC 903
Qy 450 GlyLysLeuThrGluLysSerAspValPheSerPheGlyValValLeuLeuGluLeuIle 469
Db 904 GGTCTTCTGATGAGAAAGCGATGCTACAGCTTCGGGGTTGTACTCTTGGAAAGCTATA 963
Qy 470 ThrGlyArgArgProValAspAlaAenAenValTyrValAspAspSerLeuValAspTyr 489
Db 964 ACTGATAGATATCCGTAGACTATGCTCGTCCACCACCGAGGTACATTGGTGGAGTGG 1023
Qy 490 AlaArgProLeuLeuAenArgAlaSerGluGlnGlyAspPheGlyGlyLeuAlaAspAla 509
Db 1024 CTGAAGATGATGTC-----CAACAAAGACGATCAGAAGAAAGTGGTTGATCCA 1071
Qy 510 LysMetAenAenGlyTyrAspArgGluGluMetAlaArgMetValAlaCysAlaAlaAla 529
Db 1072 AACCTTGAACAAACCACTTACAAAGTCTTGAAGAACACACTATTGACTGCTTTCAGA 1131
Qy 530 CysValArgHisSerAlaArgArgProArgMetSerGlnIleValArgAlaLeuGlu 549
Db 1132 TGTGTTGATCCAAATGCTGAGAAAGACCGGAGTATGACCAAGTTGACGATGCTTGA 1191
```

RESULT 10

```
US-09-228-986-11
; Sequence 11, Application US/09228986
; Patent No. 6359198
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Niels
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signalling
```

```
FILE REFERENCE: 11000/1020
CURRENT APPLICATION NUMBER: US/09/228,986
CURRENT FILING DATE: 1999-01-12
NUMBER OF SEQ ID NOS: 130
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 11
LENGTH: 3097
TYPE: DNA
ORGANISM: Eucalyptus grandis
US-09-228-986-11

Alignment Scores:
Pred. No.: 4,95e-17 Length: 3097
Score: 600.50 Matches: 173
Percent Similarity: 46.08% Conservative: 80
Best Local Similarity: 31.51% Mismatches: 172
Query Match: 17.39% Indels: 126
DB: 4 Gaps: 13
```

US-10-086-464-2 (1-647) x US-09-228-986-11 (1-3097)

```
Qy 66 ProSerThrProGlySerProProPro-----LeuPro 76
Db 1523 CCTGCAAGTTCTGCGACGCTCCAAAACACAGCTTAGACGGAGTTGTTCCCGACGCTTGG 1582
Qy 77 GlnProSerProAlaProThrProGlySerProProAlaProValThrPro--- 95
Db 1583 AGAGCTGAAAAGACCTTCACCTACTGACCTTGGAGATAAAGCTGTCTACAAG-GTACCTTAC 1641
Qy 96 -----ProThrArgAenProProSerValProGlyProPro 108
Db 1642 CAGATTCTTTGAAACAGACAGAGCTTGGAGGTCAGAACCTCAGGAAACTTGTGCTTTCTCT 1701
Qy 109 SerAenProSerArgGluGlyGlySerProArgProProSerSerProSerPro--- 126
Db 1702 TCTCCACAACCGCGTGC-GGTGATGCATCATCTAGTCTTCAATTGAGGCACCCCAAGTT 1760.
Qy 127 -----ProSerProSerSerAspGlyLeuSerThrGlyValValValGlyIleAla 143
Db 1761 ACAATAGTTCCCGAGAGAAACAAGGGGACATATCGTTAGCATATATCTCGGAGCA 1820
Qy 144 IleGlyGlyValAlaLeuLeuValIleValThrLeuIleCysLeuLeuCysLysLys 163
Db 1821 GTCGGAGGAGTATCACTAGCTATTTTACTCATCCGCTCTTCGTATTTCATGTACAGAGG 1880
Qy 164 ArgArgArgAspGluAspAlaTyrTyrValProProProProProGlyProLys 183
Db 1881 AGAGGAAGAACTGAA----- 1895
Qy 184 AlaGlyGlyProTyrGlyGlyGlnGlnGlnTrpArgGlnGlnAenAlaThrProPro 203
Db 1895 ----- 1895
Qy 204 SerAspHisValValThrSerLeuProProProLysAlaProSerProProArgGln 223
Db 1895 ----- 1895
Qy 224 ProProProProProProPheMetSerSerSerGlySerAspTyrSerAsp 243
Db 1896 -----ATGTCATATACGAA 1910
Qy 244 ArgProValLeuProProProSerProGlyLeuValLeuGlyPheSerLysSerThrPhe 263
Db 1911 AGGCGAGTCCGACAGCTGAGAAACTGGAAT-----GCAGTAAGATTTT 1955
Qy 264 ThrTyrGluGluLeuAlaArgAlaThrAsnGlyPheSerGluAlaAenLeuLeuGlyGln 283
Db 1956 TCCTACAAGAGATCAAAACAGCTCAACAACACTTTAAAGAA-----GTCAATGGTCAT 2009
Qy 284 GlyGlyPheGlyTyrValHisLysGlyValLeuProSerGlyLysGluValAlaValLys 303
Db 2010 GGAAGTTTGGATCCGCTGTACTTCCAGTGGGAAACCTTCCAGTTGGAACACTAGTTGCTGTGAAA 2069
```


Qy	128	rProSerSerAspGlyLeuSerThrGlyValValValGlyIleAlaIleGlyGlyValAl	148
Dd	520	TCCAGGCTCCCAGACCAATTC-	541
Qy	148	aLeuLeuValIleValThrLeuIleCysLeuLeuCysIlysIysIysArgArGArgAspGl	168
Dd	541	-	541
Qy	168	uGluAspAlaTyrTyrValProProProProProGlyProLysAlaGlyGlyProTy	188
Dd	542	-GGGCCTCGAGCTCGGC-	556
Qy	188	rGlyGlyGlnGlnGlnTrpArgGlnGlnAsnAlaThrProProSerAspHisValVa	208
Dd	556	-	556
Qy	208	lThrSerLeuProProProLysAlaProSerProProArgGlnProProProProPr	228
Dd	557	-CTGGTTCCAAGCCCTGCTTCCCTGTGGCCCTCCACC GCC	594
Qy	228	oProProPheMetSerSerSerGlyGlySerAspTyrSerAspArgProValLeuPr	248
Dd	595	ATCTCACCCCTTCTCTACCAAGCAGGC-	627
Qy	248	oProProSerProGlyLeuValLeuGlyPheSerLysSerThrPheThrTyrGlu-	266
Dd	628	AGAGAGCTCAGTGCTCTCTGCAGGAGCCGCCCTCTCCGTTTTGCTGGCCCCCTGTG	687
Qy	267	GluLeuAlaArgAlaThrAanGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyGlyPh	286
Dd	688	TGAGATTTCGGGGGACCCACAATTCTCGGAGAGCTCAAGATCGGGAGGTGGCTT	747
Qy	286	eGlyTyrValHisIysGlyValLeuProSerGlyLysGluValAlaValIysGlnLeuLy	306
Dd	748	TGGTGTGGTGTAACGGGGCGGTGATG--AGGAACACGGTGATGCTGTGAAGAGCGCTGA	804
Qy	306	sValGlySerGly-----GlnGlyGluArgGluPheGlnAlaGluValGluIl	322
Dd	805	GGAAACGCTGACCTGGAGTGGACTGCAGTAGAGCAGAGCTTCCTGACCAGGTGGAGCA	864
Qy	322	eIleSerArgValHisArgHisLeuValSerLeuValGlyTyrCysIleAlaGlyAl	342
Dd	865	GCTGTCCAGGTTTCGTCAACCAACATGTGGACTTTGCTGGCTACTGTGCTCAGAACGG	924
Qy	342	aLysArgLeuLeuValTyrGluPheValProAsnAsnLeuGlu-----LeuHisLe	360
Dd	925	CTTCTACTGCTGTGTACGGCTTCTGCCCCACGGCTCCCTGGAGACCGTCTTCACTG	984
Qy	360	uHisGlyGluGlyArgProThrMetGluTrpSerThrArgLeuLysIleAlaLeuGlySe	380
Dd	985	CCAGACCCAGCGCTGCCACCTCTCTCTGGCCCTCAGCAGCTGCATCTCTTCTGGGTAC	1044
Qy	380	rAlaLysGlyLeuSerTyrLeuHisGluAspCysAsnProLysIleIleHisArgAspIl	400
Dd	1045	AGCCCGGGCAATTCAAGTTCTACATCAGGAC--AGCCCCAGCCCTCATCTCATGGAGACAT	1101
Qy	400	eLysAlaSerAsnIleLeuIleAspPheLysPheGluAlaLysValAlaAspPheGlyLe	420
Dd	1102	CAAGATGTCACAGCTCTCTCGATGAGAGCTGCACCCCAAGCTGGAGACTTTGGCCT	1161
Qy	420	uAlaLysIle-----AlaSerAspThrAsnThrHisValSerThrArg--	434
Dd	1162	GGCCCGGTTACAGCGCTTTGCGGGTCCAGGCCCCAGCAGCAGCATGTTGGGCCCGGAC	1221
Qy	435	-----ValMetGlyThrPheGlyTyrLeuAlaProGluTyrAlaAlaSerGlyLysLe	452
Dd	1222	ACAGACAGTGGGGGACCCCTTGGCTTACTGTCGCCGAGAGGTACATCAAGACGGGAAGCT	1281
Qy	452	uThrGluLysSerAspValPheSerPheGlyValValLeuLeuLeuIleThrGlyAr	472
Dd	1282	GGCTGTGGACCGGACACCTTCACTTTGGGGTGGTAGTGCTAGAGACCTTGGCTGTGCTCA	1341
Qy	472	gArgProVal-----AspAlaAsnAsnValTyrValAspAspSerLeuValAspTr	489

```

1342 GAGGGCTGTGAAGACGACCGGTGCCAGGACCAAGTATCTGAAAGAC---CTGGTGGAA-- 1399
489 pAIAaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGluGlyLeuAlaAspAl 509
1397 -----GAGGAGCTGAGAGGCTGGAGTGGCTTTTGAGAGACACCAGAG 1440
509 aLYsMetAsnAsnGlyTyAspArgGluGluMetAlaArgMetValala----- 525
1441 CACACTGCAAGCAGGTCTGGCTGCAGATGCTGGGCTGCTCCATCGCCATGCAGATCTA 1500
525 ----- 525
1501 CAAGAAGCACCTGGACCCGAGCCGCGGCGCTGCCACCTGAGCTGGGCTGGGCTGGG 1560
526 ---CysAlaAlaAlaCysValArgHisSerAlaArgArgProArgMetSerGlnIl 544
1561 CCAGCTGGGCTGTCTGTCTGTGCACCGCGGGCCAAAGAGAGGCTCTCTATGACCAGGT 1620
544 eValhArgAlaLeuGluGlyAsnValSerLeuSerAspLeuAsnGluGlyMetArgProGl 564
1621 GTACGAGAGCTAGAG-----AAGCTGCAGGCAGTGGTGGCGGGGGTG---CCCGG 1668
564 Y-----GlnSerAsnValTySerSer-- 571
1669 GCATTTGGAGCGCGCCAGCTGCATCCCGCTTCCCGCAGGAGAACTCCTACGTGTCAG 1728
572 -----TyrgLYsSerThrAspTyAsp----- 579
1729 CACTGGCAGAGCCCAcAGTGGGGCTCTCCATGGCAGCGCCCTGGCGAGCCCATCAGGAGC 1788
580 -SerSerGlnTyAsnGluAspMetLYsPheArgLYsMetAlaLeuGlyThrGlnGl 599
1789 CAGTGGCCAGGCAGAGCAGCTGCAGAGAGGCCCAACACAGCCGCTGGAGAGTGACGA 1848
599 uTyAsnAlaThrGlyGluTySerAsnProThrSerAspTyrgLYsLeuTyProSer-- 618
1849 G-----AGCTAGGGCGGCTCTCTGCTCGCTCGGCTCTGGCACTTGACTCCAAGCTG 1902
619 -----GlySerSerSerGluGlyGlnThrThrArgGl 629
1903 CCCTCTGGACCCAGCACCCCTCAGGAGGCGGGCTGTCTCCTAGGGGAGCACGGC-AGGAG 1961
629 uMetGluMetGly-----LYsIleLYsArgThrGlyGlnGlyTySerGlyPro 645
1962 AATCGAGCTGGGGAGTGGCCAGGATCCGGGCCACACGCCCTGGAAGGACTGGCCCT 2019

RESULT 12
US-09-016-434-1093
; Sequence 1093, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,434
; FILING DATE: HEREWITH
; CLASSIFICATION:

```


1501	1441	CA	CACTGCACGACGCTCGCTGCAGATGCTCGGGCTGCTCCCATGCCATGACGATCTA	1501
525	525	---	-----	525
1501	1501	CA	AGAAGCACCTGGACCCAGCCGGCCCTGCTCCACCTGAGCTGGGCTGGGCGCTGGG	1560
526	526	---	-----CysAlaAlaCysValArgHisSerAlaArgArgProArgMetSerGlnI	544
1561	1561	CC	AGCTGGCTGCTGCTGCTGCACCCCGGGCCAAAGGAGGGCTCTCTATGACCCAGGT	1620
544	544	eva	lAtgAlaLeuGluGlyAsnValSerAspLeuAsnGluGlyMetArgProG	564
1621	1621	GT	ACGAGGCTAGAG-----AAGCTGCAGGCGAGTGGTGGCGGGGTG---CCCGG	1668
564	564	Y	-----GlnSerAsnValTyrSerSer--	571
1669	1669	GC	ATTTGGAGGCGCCAGCTGCATCCCCCTTCCCGCAGGAGACTCCTACGTGTCCAG	1728
572	572	---	-----TyrGlyGlySerThrAspTyrAsp-	579
1729	1729	CA	CTGGCAGAGCCACAGTGGGGCTGTCTCATGTGCAGCCCTGGCAGCGCATCAGGAGC	1788
580	580	-Ser	SerGlnTyrAsnGluAspMetLysLysPheArgLysMetAlaLeuGlyThrGlnG	599
1789	1789	CA	TGCCCGCCAGCAGCAGCAGCTGCAGAGAGGCCCCAACACAGCCCGTGGAGAGTGACGA	1848
599	599	uTyr	AsnAlaThrGlyGluTyrSerAsnProThrSerAspTyrGlyLeuTyrProSer--	618
1849	1849	G	-----AGCTAGGCGGCTCTCTGTGTCGCTCGCGCTCTGGCCTTGACTTGCACCTG	1902
619	619	---	-----GlySerSerSerGlyGlyGlnThrThrArgG	629
1903	1903	CC	CTCTGGACCCAGCACCCCTCAGGGAGGCGGCTGTCTCAGGGGACACGGC-AGGAG	1961
629	629	uMet	GluMetGly-----LysIleLysArgThrGlyGlnGlyTyrSerGlyPro	645
1962	1962	AA	TGAGCTGGGGAGTGGCCAGGATCTCCGGCCCCACAGCCGTGGAGAGGACTGGCCCT	2019
RESULT 13				
PCT-US96-09193-1				
; Sequence 1, Application PC/TUS9609193				
; GENERAL INFORMATION:				
; APPLICANT: Tularik, Inc.				
; TITLE OF INVENTION: INTERLEUKIN-1 RECEPTOR-ASSOCIATED				
; TITLE OF INVENTION: PROTEIN KINASE AND BINDING ASSAY				
; NUMBER OF SEQUENCES: 2				
; CORRESPONDENCE ADDRESS:				
; ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT				
; STREET: 4 Embarcadero Center, Suite 3400				
; CITY: San Francisco				
; STATE: California				
; COUNTRY: USA				
; ZIP: 94111-4187				
; COMPUTER READABLE FORM:				
; MEDIUM TYPE: Floppy disk				
; COMPUTER: IBM PC compatible				
; OPERATING SYSTEM: PC-DOS/MS-DOS				
; SOFTWARE: Patent In Release #1.0, Version #1.30				
; CURRENT APPLICATION DATA:				
; APPLICATION NUMBER: PCT/US96/09193				
; FILING DATE: JUNE 5 1996				
; CLASSIFICATION:				
; PRIOR APPLICATION DATA:				
; APPLICATION NUMBER: U.S. Serial No. 08/587,889				
; FILING DATE: JAN 16 1996				
; CLASSIFICATION:				
; APPLICATION NUMBER: U.S. Serial No. 08/494,006				
; FILING DATE: JUNE 23 1995				
; CLASSIFICATION:				
; ATTORNEY/AGENT INFORMATION:				
; NAME: David J. Brezner				
; REGISTRATION NUMBER: 24,774				
; REFERENCE/DOCKET NUMBER: FP-62191-1				

Db 557 -----CTGGTTCCAGCCCTCTTCCCTGTGGCTCCACCGCC 594
Qy 228 oProProPhoMetSerSerGlyGlySerAspTyrSerAspArgProValLeuPr 248
Db 595 ATCTCCAGCCCTCTTCTTACCAAGCCAGGC-----CC 627
Qy 248 oProProSerProGlyLeuValLeuGlyPheSerIysSerThrPheThrTyrGlu----- 266
Db 628 AGAGAGCTCAGTGTCCCTCTCGAGGGAGCCGCCCTCTCCGTTTCTGCTGCCCTCTG 687
Qy 267 -GluLeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyGlyPh 286
Db 688 TGAGATTCCCGGGCACCACCACTTCTCGAGAGCTCAAGATCGGGAGGTGGCTT 747
Qy 286 eGlyTyrValHisIstysGlyValLeuProSerGlyIysGluValAlaValIysGlnLeuLy 306
Db 748 TGGGTGCTGTACCGGGCGGTGATG--AGAAACACGGGTGTATGCTGTGAAGAGCTGAA 804
Qy 306 sValGlySerGly-----GlnGlyGluArgGluPheGlnAlaGluValGluI 322
Db 805 GGAGAACCTGACCTGGAGTGCAGTGAAGCAGAGCTTCTCGACCGAGGTGGAGCA 864
Qy 322 eIleSerArgValHisHisArgHisLeuValSerLeuValGlyTyrCysIleAlaGlyAl 342
Db 865 GCTGTCCAGGTTTCGTACCCCAACATTTGTGGACTTTCCTGGCTACTGTGCTCAGAACGG 924
Qy 342 aLysArgLeuLeuValTyrGluPheValProAsnAsnAsnLeuGlu-----LeuHisLe 360
Db 925 CTTCTACTGCTGTGTGCGCTTCTGCCCAACGGCTCCCTGGAGGACCGCTCTCCACTG 984
Qy 360 uHisGlyGluGlyArgProThrMetGluTyrSerThrArgLeuIysIleAlaLeuGlySe 380
Db 985 CCAGACCCAGCGCTGCCCACTCTCTCTGGCTCAGCGCTGAGACATCTCTTCTGGGTAC 1044
Qy 380 rAlaIysGlyLeuSerTyrLeuHisGluAspCysAsnProIysIleIleHisArgAspI 400
Db 1045 AGCCCGGCAATTTCAGTTTCTACATCAGAC---AGCCCCAGCTCATCTCCATGGAGCAT 1101
Qy 400 eLysAlaSerAsnIleLeuIleAspPheGlyPheGluAlaIysValAlaAspPheGlyLe 420
Db 1102 CAAGAGTTCCAAACGTCTCTGTGATGAGAGGTGACACCAAGCTGGGAGACTTTGGCCT 1161
Qy 420 uAlaIysIle-----AlaSerAspThrAsnThrHisValSerThrArg-- 434
Db 1162 GCGCGGTTACCGCTTTTGGCGGTCCAGCGCCAGCCAGCAGCATGTTGGCGCCGAC 1221
Qy 435 -----ValMetGlyThrPheGlyTyrLeuAlaProGluTyrAlaAlaSerGlyLysLe 452
Db 1222 ACAGACATGCGGGCACCCTGGCTTACTGCCGAGGAGTACATCAAGACGGGAGGCT 1281
Qy 452 uThrGluIysSerAspValPheSerGlyValValLeuLeuGluLeuIleThrGlyAr 472
Db 1282 GGCTGTGACACGACACCTTTCAGCTTTGGGTGGTGTAGTGTAGAGACCTTGGCTGTCA 1341
Qy 472 gArgProVal-----AspAlaAsnAsnValTyrValAspAspSerLeuValAspTr 489
Db 1342 GAGGCTGTGAAGACGACCGTGCAGATATCTGAAAGAC---CTGGTGA-- 1396
Qy 489 pAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGluGlyLeuAlaAspAl 509
Db 1397 -----GAGGAGGTGTAGAGGCTGGAGTGGCTTGTGAGAGCACCACCGAG 1440
Qy 509 aLysMetAsnAsnGlyTyrAspArgGluGluMetAlaArgMetValAla----- 525
Db 1441 CACACTGCAAGCAGGTCTGGCTGCAGATCTGGGCTGCTCCCATCGCCATGCAGATCTA 1500
Qy 525 ----- 525
Db 1501 CAAGAACGACCTGGACCCCGCGGCCCTGCCACCTGAGCTGGGCTGGCGCTGGG 1560
Qy 526 ----CysAlaAlaAlaCysValArgHisSerAlaArgArgProArgMetSerGlnI 544
Db 1561 CCAGCTGGCTGCTGCTGCTGCACCGCGGGCCCAAGAGGCGCTCTCTATGACCCAGGT 1620

Qy 544 eValArgAlaLeuGlnGlyAsnValSerLeuSerAspLeuAsnGluGlyMetArgProGl 564
Db 1621 GTACGAGAGGCTAGAG-----AAAGCTGACGAGCAGTGGTGGCGGGGTG---CCCGG 1668
Qy 564 Y-----GlnSerAsnValTyrSerSer-- 571
Db 1669 GCATTTGGAGCCGCCAGCTGCATCCCTTCCCGCAGGAGAACTTCTACGTGTCCAG 1728
Qy 572 -----TyrGlySerThrAspTyrAsp----- 579
Db 1729 CACTGGCAGAGCCACACAGTGGGCTGTCTCCATGGCAGCCCTGGCAGCGCCATCAGGAGC 1788
Qy 580 -SerSerGlnTyrAsnGluAspMetLysLysPheArgLysMetAlaLeuGlyThrGlnGl 599
Db 1789 CAGTGGCCAGCAGCAGCAGCAGCTGCAGAGAGGGCCCAACACGAGCCGCTGGAGAGTGA 1848
Qy 599 uTyrAsnAlaThrGlyGluTyrSerAsnProThrSerAspTyrGlyLeuTyrProSer-- 618
Db 1849 G-----AGCTAGCGGCTCTCTGTGCGCTCGGCTCTGGCACTTGACTTCCAGCTG 1902
Qy 619 -----GlySerSerSerGluGlyGlnThrThrArgGl 629
Db 1903 CCCTCTGGACCCAGCACCCTCAGGAGGCGCGGTCTCTCAGGGGACACGGC-AGGAG 1961
Qy 629 uMetGluMetGly-----LysIleLysArgThrGlyGlnGlyTyrSerGlyPro 645
Db 1962 AATCAGCTGGGGAGTGGCCAGGATCCCGGCCACAGCCGCTGGAAGGACTGGCCCT 2019
RESULT 14
US-09-602-472A-5
; Sequence 5, Application US/09602472A
; Patent No. 6608240
; GENERAL INFORMATION:
; APPLICANT: Bidney, Dennis L.
; APPLICANT: Hu, Xu
; APPLICANT: Li, Guihua
; TITLE OF INVENTION: Sunflower Disease Resistance Genes
; FILE REFERENCE: 35718/200630
; CURRENT APPLICATION NUMBER: US/09/602,472A
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/140,876
; PRIOR FILING DATE: 1999-06-24
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 2114
; TYPE: DNA
; ORGANISM: Helianthus annuus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1) ... (2114)
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1475
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc_feature
; LOCATION: 1475
; OTHER INFORMATION: n = A,T,C or G
US-09-602-472A-5
Alignment Scores:
Pred. No.: 2,96e-16 Length: 2114
Score: 578.00 Matches: 143
Percent Similarity: 55.07% Conservative: 85
Best Local Similarity: 34.54% Mismatches: 136
Query Match: 16.74% Indels: 51
DB: 4 Gaps: 15
US-10-086-464-2 (1-647) x US-09-602-472A-5 (1-2114)

Qy	235	SerSerGlyGlySerAspTyrSerAspArgProValLeuProProSerProGlyLeu	254
Db	52	GCTGAATCTCTCTCATATTTCCACCAACCGTCG	87
Qy	255	ValLeuGlyPheSerLysSerThrPheThrTyrGluGluLeuAlaArgAlaThrAsnGly	274
Db	88	-----CGTCAATTATACCTTTTCGGAGATTCAACTTGCACACCCAAAC	129
Qy	275	PheSerGluAlaAsnLeuLeuGlyGlnGlyPheGlyTyrValHisLysGlyValLeu	294
Db	130	TTTGATGAGTCGTGGTGATAGGACGCTGGGGGTTTGGCAAGGTTTACAGAGGAACCTTC	189
Qy	295	ProSerGlyLys-----GluValAlaValLysGlnLeuLysValGlySerGlyGln	311
Db	190	ACTTATGGGGAACATACCTGGATGCTGCAATTAAGCGACTGGATCAGGTTCTTAGTCAA	249
Qy	312	GlyGluArgGluPheGlnAlaGluValGluLeuLeuSerArgValHisArgHisLeu	331
Db	250	GGACGGTAGAGTTTAGGCGCTGAATAGAGATGCTCTCAATCTAAGGCATCTGTCATTG	308
Qy	332	ValSerLeuValGlyTyrCysIleAlaGlyAlaLysArgLeuLeuValTyrGluPheVal	351
Db	310	GTGCTTTTAATGGTACTAGTAGTGATGGCGCAAGAGATGGTCTTGATATGAACATATG	369
Qy	352	ProAsnAsnLeuLeuLeuHisLysHisGlyGluGlyArgProThrMetGluTrpSer	371
Db	370	CCCAATGGAACCTTTGCAGATCGCTCCAC---AAGCGTCAGAGCTCCTCTAACTTGGGTA	426
Qy	372	ThrArgLeuLysIleAlaLeuGlySerAlaLysGlyLysSerTyrLeuHis-----Glu	389
Db	427	AGAAGACTCAAAATATGCATAGGGCGCGCTCGTGGTTTAGATTACTTTGCACACTGGTACG	486
Qy	390	AspCysAsnProLysIleHisArgAspIleLysAlaSerAsnIleLeuIleAspPhe	409
Db	487	GGTATTAACCATGGAGTTATACATCGGATCTTAAGAGCAAAATATATTGTTAGATCAC	546
Qy	410	LysPheGluAlaLysValAlaAspPheGlyLeuAlaLysIle-----AlaSerAspThr	427
Db	547	AAATGGGCAGCTAAGGTTTCTGACTTTGGTTTGTCCAAAATTGGTCCAAACAATCAGCGCT	606
Qy	428	AsnThrHisValSerThrArgValMetGlyThrPheGlyTyrLeuAlaProGluTyrAla	447
Db	607	TCAACTATTGTATACACTTTGGTGAGAGGCACCTTTGGATATATGGATCCAGATTACTTT	666
Qy	448	AlaSerGlyLysLeuThrGluLysSerAspValPheSerPheGlyValValLeuLeuGlu	467
Db	667	CAACAGGTAGGCTGACTCGAAAGCTGACGTGTATGCCCTTTGGGGTGGTCTCTTTGAA	726
Qy	468	LeuIleThrGlyArgArgProValAspAlaAsnAsnValTyrValAspSerLeuVal	487
Db	727	GTCTCTGTGGGAACAAGTAGTAGTATGATGAGGAGCAC-TGG-----GGTTTGCA	776
Qy	488	AspTrpAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGluGlyLeuAla	507
Db	777	ACATGGGCTCAAGACTCTCTT-----AAAGAAGGAAGCGCTAAAGCAAAATTTGTT	824
Qy	508	AspAlaLysMetAsnAsnGlyTyrAspArgGluGluMetAlaArgMetValAlaCysAla	527
Db	825	GATTCTTAATTAAGGGGAGAAATATCCCAAAATGTTTGAAGAGTTTGCACACTACTAGCT	884
Qy	528	AlaAlaCysValArgHisSerAlaArgArgProArgMetSerGlnIleValArgAla	547
Db	885	GACCGGTGTTTGATAGCGCTCCCAAGCAACGTCTCGAATGGCGGAGGTTGTGATGGT	944
Qy	548	LeuGluGlyAsnValSerLeuSerAspLeuAsnGluGly-----Met	561
Db	945	CTTGAGTCAATCTTAGCGCTTACAGAGAGAAAACCGAGAGTACATGGGTACCAACATTTCTC	1004
Qy	562	ArgProGlyGlnSerAsnValTyrSerSerTyrGlyGlySerThrAspTyrAspSerSer	581
Db	1005	TTTCCA-----TTCACTTGGGAAAAGACAGTTGGAGGTACAGATTTTAAATCCCTGT	1055
Qy	582	Gln-----TyrAsn-----GluAsp-----MetLysLysPheArgLysMet	593

US-08-265-628-1

Alignment Scores:

Pred. No.:	5,04e-16	Length:	2749
Score:	575.00	Matches:	154
Percent Similarity:	49.80%	Conservative:	94
Best Local Similarity:	30.92%	Mismatches:	160
Query Match:	16.65%	Indels:	91
DB:	1	Gaps:	14

US-10-086-464-2 (1-647) x US-08-265-628-1 (1-2749)

Qy	130	SerSerAspGlyLeuSerThrGlyValValValGlyValLeuAlaGlyValAlaLeu	149
Db	1327	AACGGAATGGGAAACCATATGCTGATTTGGAGTTTGGTTT-----	1371
Qy	150	LeuValIleValThrLeuIleCysLeuLeuCysIleValLeuValArgArgAspGlu	169
Db	1372	CTGCTTCTTATGATCATGTTCTGCG---CTCTGGAAAGGAAACAAACGAGCA	1428
Qy	170	AspAlaTyrTyrValProProProProGlyProLysAlaGlyProTyrGly	189
Db	1429	ACTGCAACATCTATTGTAAT-----	1449
Qy	190	GlyGlnGlnGlnInTrpArgGlnGlnAsnAlaThrProProSerAspHisValThr	209
Db	1450	-----CGACAGAGAAAC-----CAGATTGCTAATGAAC	1479
Qy	210	SerLeuProProProProLysAlaProSerProProArgGlnProProProPro	229
Db	1480	GGGATG-----ATACTACAGCAAGACAGACATTCGCT-----	1512
Qy	230	ProProPheMetSerSerSerGlySerAspTyrSerAspArgProValLeuPro	249
Db	1513	-----ATAGAGAACAAACTCGAGGAATTGGAACITTCATTGATA-----	1551
Qy	250	ProSerProGlyLeuValLeuGlyPheSerLysSerThrPheThrTyrGluGluLeu	269
Db	1552	-----GAGTTGGAAAGCTGTGTCTC	1569
Qy	270	ArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyPheGlyTyrVal	289
Db	1570	AAAGCCACCGAAAATTTCTCCAAATGTAAACAACTCGGACACAGGTGGTTTGGTATGTT	1629
Qy	290	HisLysGlyValLeuProSerGlyLysGluValAlaValLysGlnLeuLysValGlySer	309
Db	1630	TACAAGGGTAGATTACTTGGTGGCAAGAAATGGCGGTAAAGGCTATCAAAAACGTCG	1689
Qy	310	GlyGlnGlyGluArgGluPheGlnAlaGluValGluIleIleSerArgValHisArg	329
Db	1690	GTTCAGGGGACTGGTGAGTTTATGAATGAGGTGAGATTGTCGCGAGGCTTCAGCATATA	1749
Qy	330	HisLeuValSerLeuValGlyTyrCysIleAlaGlyAlaLysArgLeuValTyrGlu	349
Db	1750	AACCTTGTCGGAATTTCTGGCTGCTGATTCAGTGGGACAGAGAAATGCTGTATATGAG	1809
Qy	350	PheValProAsnAsnLeuGluHisLeuHisGlyGluGlyArg---ProThrMet	368
Db	1810	TATTAGAAAATTTAAGCTCGATCTTATCTCTTCGGAATTAACGAGCTCTACGTTA	1869
Qy	369	GluTrpSerThrArgLeuLysIleAlaLeuGlySerAlaLysGlyLeuSerTyrLeuHis	388
Db	1870	AATTGGAAGGACAGATTCAACATTACCAATGGTGTGCTCGAGGACTTTTATATCTTCAT	1929
Qy	389	GluAspCysAsnProLysIleIleHisArgAspIleLysAlaSerAsnIleLeuIleAsp	408
Db	1930	CAAGACTCAGGTTTAGGATATCCACAGAGATATGAAGTAAGTAACATTTGCTGAT	1989
Qy	409	PheLysPheGluAlaLysValAlaAspPheGlyLeuAlaLysIleAlaSerAspThrAsn	428
Db	1990	AAAAATATGACACCAAGATCTCGGATTTGGGATGGCCAGAAATCTTTGCAAGGACGAG	2049
Qy	429	ThrHisValSerThrArg---ValMetGlyThrPheGlyTyrLeuAlaProGluTyrAla	447

Search completed: April 25, 2004, 01:02:38

Job time : 181 secs

Db	2050	ACTGAAGCTAACACAGGAAGGTGGTCGAACTTACGGCTACATGTCTCCGAGTAGCCCA	2109
Qy	448	AlaSerGlyLysLeuThrGluLysSerAspValPheSerPheGlyValValLeuGlu	467
Db	2110	ATGGATGGGGTATTCTCGGAAAAATCAGATGTTTTCAGTTTGGAGTCATTGTTCTTCAA	2169
Qy	468	LeuIleThrGlyArgArgProValAspAlaAsnAsnValTyrValAspAspSerLeuVal	487
Db	2170	ATTGTTAGTGGAAAAAGGACAGAGGATTCTACAACTTGAACCAACGAAAAAATCTCTTA	2229
Qy	488	Asp-----TrpAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAsp-----	502
Db	2230	AGCTATGATGAGTCTACTGGACGAGGAAAGCGCTAGAAATTTGTATCCAGTCATC	2289
Qy	503	PheGluGlyLeuAlaAspAlaLysMetAsnAsnGlyTyrAspArgGluGluMetAlaArg	522
Db	2290	GTAGATTCTATTGTCATCTATTACCAGCAACC-----TTTCAACCAAAAAGATTCTATAA	2343
Qy	523	MetValAlaCysAlaAlaAlaCysValArgHisSerAlaArgArgProArgMetSer	542
Db	2344	TGCATACAAATTGGTCTCTTGTGTTCAGAACGTCGACAGCATAGACCAACGATGTCG	2403
Qy	543	GlnIleValArgAlaLeuGluGlyAsnValSerLeuSerAspLeuAsnGluGlyMetArg	562
Db	2404	TCCGTGGTGGATGCTTGGAAAGTGAAGCA-----ACGAGATTCCTGAGCCTACACCG	2457
Qy	563	ProGlyGlnSerAsnValTyrSerSerTyrGlyGlySer-----	575
Db	2458	CCAGTTATTCCCTCGGAAGAAGTCCCTTATGAAAAATAATCTTCATCAAGTAGACATTGC	2517
Qy	576	-----ThrAspTyrAspSerSerGln	582
Db	2518	GACGACGACGAATCCTGGACGCTGAACCCAGTACACCTGCTCAGACATCGATCCCGGTAG	2577
Qy	583	TyrAsnGluAspMetLysLysPheArgLysMetAlaLeuGlyThrGlnGluTyr	600
Db	2578	TACGA-AATCCGTTGAGAAAGTTTCAGATAATTAACTATTGGGGTGACCCGGATAT	2630

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 25, 2004, 00:01:06 ; Search time 638 Seconds
(without alignments)
4572.228 Million cell updates/sec

Title: US-10-086-464-2

Perfect score: 3453

Sequence: 1 MSSAPSGTGPSPSPNST.....REMEGKIKRTGGQSGPSL 647

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2907579 seqs, 2254313464 residues

Total number of hits satisfying chosen parameters: 5815158

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10086464/runat_23042004_083017_5726/app_query.fasta_1.839.
-DB=Published Applications NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bites -START=1 -END=1 -MATRIX=blousem62
-TRANS=human40.cdi -LIST=45 -DLOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTPMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10086464 @CGN 1 1 333 @runat 23042004 083017_5726
-NCFU=6 -ICFU=3 -NO_WMAP -LARG_QUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARG_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
------------	-------	-------------	--------	----	-------------

RESULT 1
US-10-086-464-1
; Sequence 1, Application US/10086464
; Publication No. US20020199218A1
; GENERAL INFORMATION:
; APPLICANT: GORING, Daphne R. et al.
; TITLE OF INVENTION: PROLINE-RICH EXTENSIN-LIKE RECEPTOR KINASES
; FILE REFERENCE: P 25,762-A USA
; CURRENT APPLICATION NUMBER: US/10/086,464
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 10/069,304
; PRIOR FILING DATE: 2002-02-19
; PRIOR APPLICATION NUMBER: PCT/CA00/00966
; PRIOR FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: US 60/149,466
; PRIOR FILING DATE: 1999-08-19
; PRIOR APPLICATION NUMBER: US 60/159,422
; PRIOR FILING DATE: 1999-10-13
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1944
; TYPE: DNA

ALIGNMENTS


```
Db 1921 GGTATAGTGACCTTCTT 1941
RESULT 2
US-10-086-464-3
; Sequence 3, Application US/10086464
; Publication No. US20020199218A1
; GENERAL INFORMATION:
; APPLICANT: GORING, Daphne R. et al.
; TITLE OF INVENTION: PROLINE-RICH EXTENSIN-LIKE RECEPTOR KINASES
; FILE REFERENCE: P 25,762-A USA
; CURRENT APPLICATION NUMBER: US/10/086,464
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 10/069,304
; PRIOR FILING DATE: 2002-02-19
; PRIOR APPLICATION NUMBER: PCT/CA00/00966
; PRIOR FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: US 60/149,466
; PRIOR FILING DATE: 1999-08-19
; PRIOR APPLICATION NUMBER: US 60/159,122
; PRIOR FILING DATE: 1999-10-13
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 2189
; TYPE: DNA
; ORGANISM: Brassica napus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2189)
; OTHER INFORMATION:
US-10-086-464-3
Alignment Scores:
Pred. No.: 1,02e-215 Length: 2189
Score: 3453.00 Matches: 647
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0
US-10-086-464-2 (1-647) x US-10-086-464-3 (1-2189)
Qy 1 MetSerSerAlaProSerProGlyThrGlySerProProSerProSerAsnSerThr 20
Db 97 ATGTCCTCGGCGCGTCTCGGGGACTGGTTCGCTCCATCTCCACCATCAAACTCCACA 156
Qy 21 ThrThrThrProProAlaSerAlaProProThrThrProSerSerProProPro 40
Db 157 ACCACCACTCTCTCCAGCTTCGGCTCTCTCCCTCCACCAACACACCTTCTCTCCGCGC 216
Qy 41 ProSerThrThrProSerProProSerProProSerSerArgSerThrProSerAlaProPro 60
Db 217 CCATCCACTATTCGACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 276
Qy 61 ProSerProProThrProSerThrProGlySerProProProProProProProSerPro 80
Db 277 CCATCTCCACCACTCCATCTACGCGGGATCTCCACCTCTCTCTCTCTCTCTCTCTCTCTCCA 336
Qy 81 ProAlaProThrThrProGlySerProProAlaProValThrProProThrArgAsnPro 100
Db 337 CCCGCTCCAACTACGCGCGGATCTCCACCCGCACTGTGTACTCTCTCTCTCTCTCTCTCTCT 396
Qy 101 ProProSerValProGlyProProProSerAsnProSerArgGluGlyGlySerProArgPro 120
Db 397 CCACCTTCAGTCCCGACCAACCGTCCAATCTTCACGCGAAGGAGGATCTCTCTCGACCT 456
Qy 121 ProSerSerProSerProProSerProSerSerGlyLeuSerThrGlyValValVal 140
Db 457 CCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 516
Qy 141 GlyLeuAlaLeuGlyValAlaLeuValThrLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 160
Db 517 GGAATCGCATCGGAGGAGTGGCTCTGCTTGTGATAGTACTCTGATTTGTCTCTCTCTCTCTCT 576
Qy 161 LysLysLysArgArgArgAspGluGluAspAlaTyrTyrValProProProProProPro 180
Db 577 AAGAGAAACGACGAGAGAGAGAGATGCTTACTATGTCTCTCTCTCTCTCTCTCTCTCTCTCT 636
Qy 181 GlyProLysAlaGlyGlyProTyrGlyGlyGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200
Db 637 GGTCCCAAGCCGAGGACCTTACGGTGGAGCAGCAGCAACATGGCGGCAACAACAGCA 696
Qy 201 ThrProProSerAspHisValValThrSerLeuProProProProProProProProProPro 220
Db 697 ACACCCCGTCAGATCATGTCGTGACGTCTACTACCACCACTTAAGGCTTCAATCTCCA 756
Qy 221 ProArgGlnProProProProProProProPheMetSerSerSerGlyGlySerAsp 240
Db 757 CCACGGCAACTCTCTCCACCTCCACCAACCGCTTTCATGACGACGACGCGGCTCCGAC 816
Qy 241 TyrSerAspArgProValLeuProProSerProGlyLeuValLeuGlyPheSerIys 260
Db 817 TACTCGGACGTCAGTCTCTCTCCACCGCTCTCCAGGGCTTGTGTAGGCTTCTCAAA 876
Qy 261 SerThrPheThrTyrGluGluLeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeu 280
Db 877 AGCACITTTACATACGAGGAGCTAGCTAGAGCCACCAATGGTTTCTCCGAGCGCACTTG 936
Qy 281 LeuGlyGlnGlyGlyPheGlyTyrValHisLysGlyValLeuProSerGlyLysGluVal 300
Db 937 TTAGGACACGCGGTTTCGGTTACGTGCACAAAGTGTGTTCCTAGTGGGAAAGAGTT 996
Qy 301 AlaValLysGlnLeuLysValGlySerGlyGlnGlyGluArgGluPheGlnAlaGluVal 320
Db 997 GCTGTGAAGCAGTTGAAAGTTGGAGTGGTTCAGGAGAGAGGAGTTTCAGGACAGAGTT 1056
Qy 321 GluLeuSerArgValHisHisArgHisLeuValSerLeuValGlyTyrCysLeuAla 340
Db 1057 GAGATCATACGACGAGTTCCACACGACATCTGGTGTCTCTTGTGTGTATTTGATCGCC 1116
Qy 341 GlyAlaLysArgLeuLeuValTyrGluPheValProAsnAsnAsnLeuGluLeuHisLeu 360
Db 1117 GGTGCCAAAGATTTGCTTGTATGAGTTTGTCTTAAACAACATCTCGAGCTTCACTC 1176
Qy 361 HisGlyGluGlyArgProThrMetGluTyrSerThrArgLeuLysIleAlaLeuGlySer 380
Db 1177 CATGGCGAGGACGCGCTACAAATGGAATGGAGCACCAGATTGAAGATTGCTCTTTGGATCT 1236
Qy 381 AlalysGlyLeuSerTyrLeuHisGluAspCysAsnProLysIleIleHisArgAspIle 400
Db 1237 GCTAAAGGACTTTCTTATCTTCATGAAGATTGCAATCTTAAATCATTTACCCGTGATC 1296
Qy 401 LysAlaSerAsnIleLeuIleAspPheGluAlaLysValAlaAspPheGlyLeu 420
Db 1297 AAGCTTCAACATATTTAGATAGATTTCAGTTTGAAGTTGCTGATTTGTGCTTT 1356
Qy 421 AlalysIleAlaSerAspThrAsnThrHisValSerThrArgValMetGlyThrPheGly 440
Db 1357 GCTAAGATTGCTTCTGATACAAACACGATGTATCAACAGTGTGATGGGAACCTTTGGG 1416
Qy 441 TyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPheSer 460
Db 1417 TACTTGGCTCCGAATACGCTGCAAGCGGAAAGCTCACGAGAGAGTCTGACGTTTCTCA 1476
Qy 461 PheGlyValValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 480
Db 1477 TTTGGCGTGTGCTTTTGGAGCTCATTTACTTGGAGCTCGACCCGTTGATGCCAACATGTC 1536
Qy 481 TyrValAspAspSerLeuValAspTyrAlaArgProLeuLeuAsnArgAlaSerGluGln 500
Db 1537 TATGTAGATCACAGCTTAGTTGACTGGGCACGACCACTTGTCTTAACCCGAGCATCTGAGCAA 1596
Qy 501 GlyAspPheGluGlyLeuAlaAspAlaLysMetAsnAsnGlyTyrAspArgGluGluMet 520
Db 1597 GGAGACTTTGAGGTTTAGCTGTGCAAGATGAATTAATGGTATGATGACAGAGAGAGATG 1656
```

```
Qy 521 AlaArgMetValAlaCysAlaAlaAlaCysValArgHisSerAlaArgArgProArg 540
Db 1657 GCTGCATGGTGTGTGCTGCGCGCTGTGTGTGCGCAATTCAGCTCGCGCAGACCTCGC 1716
Qy 541 MetSerGlnIleValArgAlaLeuGluGlyAsnValSerLeuSerAspLeuAsnGluGly 560
Db 1717 ATGAGCCAGATTGCGTGGCTAGAGGAATATGATCACTGTCAAGTCTTAACGAGGG 1776
Qy 561 MetArgProGlyGlnSerAsnValTyrSerSerTyrGlyGlySerThrAspTyrAspSer 580
Db 1777 ATGAGCCAGGTCAAAGCAATGTATACAGCTCATACGAGGAAGCACCGATTATGACTCG 1836
Qy 581 SerGlnTyrAsnGluAspMetLysLysPheArgLysMetAlaLeuGlyThrGlnGluTyr 600
Db 1837 AGCCAGTACAATGAAGACATGAAGAAGTTTAGGAAAATGGCACATTCGAACATCAAGAGTAC 1896
Qy 601 AsnAlaThrGlyGluTyrSerAsnProThrSerAspTyrGlyLeuTyrProSerGlySer 620
Db 1897 AACGCCACGGGTGAGTACAGTAATCCGACCACTGACTATGACTGATCCCGTCTGTTCA 1956
Qy 621 SerSerGluGlyGlnThrThrArgGluMetGluMetGlyLysLysLysLysArgThrGlyGln 640
Db 1957 AGCAGCGAGGCGCAAAACCAACCGGAAATGGAGATGGGAAGATTAAAGAAACCGGTACG 2016
Qy 641 GlyTyrSerGlyProSerLeu 647
Db 2017 GGTATAGTGACCTTCTCT 2037
```

RESULT 3

```
US-10-424-599-95479
; Sequence 95479, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 95479
; LENGTH: 2451
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_57229C.1
US-10-424-599-95479
```

```
Alignment Scores:
Pred. No.: 9e-140 Length: 2451
Score: 2285.50 Matches: 450
Percent Similarity: 80.86% Conservative: 74
Best Local Similarity: 89.44% Mismatches: 97
Query Match: 66.19% Indels: 27
DB: 13 Gaps: 16
```

US-10-086-464-2 (1-647) x US-10-424-599-95479 (1-2451)

```
Qy 13 ProSerProProSerAsnSerThrThrThrThrProProProAla-----SerAlaPro 30
Db 246 CCCCTCCCTCGCCCT--GCAACTCCGTCGTCAGCTCCGCGGCAACTCCGCTGCTCCT 302
Qy 31 ProProThrThrProSerSerProProProProSerThrThrThrProProProPro 50
Db 303 CCGCGCGGACACCTTCTTCCGACCTCCG-----TCAACTCCTTCTCCGCCCTCCG 356
Qy 51 SerSerArgSerProSerAlaProProProProProProProProProProProGly 70
Db 357 -----GCGACTCCCTCCGCTTCTCCACCGTCCACTCCTCTGCTTCTCCTCCACCG 407
```

```
Qy 71 SerProProProLeuProGlnProSerProProAlaProThrThrThrProGlySerProPro 90
Db 408 TCCACTCCAAACCGCGCTCAACTTCGCGCGCAATCG---ACTTCTCCGCGCTGCGCGCA 464
Qy 91 -----AlaProValThrProThrArgAsnProProProProSerValProGlyProPro 108
Db 465 TCGCACTCGCGCGCTGCGCTCCAGTGGCGCGGCGGCGGCGAGCAGCATCCGAGTCCACCG 524
Qy 109 SerAsnProSerArgGluGlyGlySerProArgProProProSerProSerProProSer 128
Db 525 TCCCGAGCTCGCCCTCTCCTCCGTCGAGATTCGAGACCGACCATCCGTCGCTCCGCTCG 584
Qy 129 ProSerSerAspGlyLeuSerThrGlyValValValGlyLysLysLysLysLysValAla 148
Db 585 TCTTCTCTCGTCGATATTCGACCGGTGTGTGTGGAAATCGCGGTGTGGGCTGTGGCG 644
Qy 149 LeuLeuValIleValThrLeuLysLeuLeuCys---LysLysLysLysLysLysLysLys 167
Db 645 GTTCTTCTGTGTGAGCATTCTCTGCATATGTTGCGGGAAGAAGAAGAAGACGTGAT 704
Qy 168 GluGluAspAlaTyrTyrValProProProProProProProProProProProProPro 186
Db 705 GAAGAG-----TACTATGCTCCGCGCGCAACCGCGCGGAGACCTAA--GATGAT 755
Qy 187 ProTyrGlyGlyGlnGlnGlnTyrArgGlnGlnAsnAlaThrProProSerAspHis 206
Db 756 GCATATGGTGGTCCCCCAGCTCAATGG---CAACACAAATGTTCCCTCTCTCAAGATCAT 812
Qy 207 ValValThrSerLeuProPro-----ProProLysAlaProSerProProArgGln 223
Db 813 GTGTCTCAATGATGCTCTCCAAAGCCATCGCCACCATCTCTCCACCGGCTTATGTG 872
Qy 224 ProProProProProProProProPheMetSerSerSerGly-----GlySerAspTyr 241
Db 873 CAACCTCCCGCACCGCGCTCTCTTTCATCAGCAGCAGTGGCGGATCTGGATCAAACTAT 932
Qy 242 SerAspArgProValLeuProProSerProGlyLeuValLeuGlyPheSerLysSer 261
Db 933 TCAGCGGTGAATTTCTCTCTCTCTCTCCAGAAATTTCTTGGGATTTCTTAAGAGC 992
Qy 262 ThrPheThrTyrGluGluLeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeu 281
Db 993 ACATTCACTATGAGGAGTTGGCACGCGCAACTGATGCTCTCTGATGCCACCTCTT 1052
Qy 282 GlyGlnGlyGlyPheGlyTyrValHisLysGlyValLeuProSerGlyLysGluValAla 301
Db 1053 GGACAGGAGGATTTGGATATGTGCACAGAGGAATTTCTTCCCAACGCGCAGGAGGTGCA 1112
Qy 302 ValLysGlnLeuLysValGlySerGlyGlnGluArgGluPheGlnAlaGluValGlu 321
Db 1113 GTGAAGCAATTGAAGGCTGGAAGCGGCAAGGAGCGTGAATTCGAAGCTGAAGTTGAG 1172
Qy 322 IleIleSerArgValHisHisArgHisLeuValSerLeuValGlyTyrCysIleAlaGly 341
Db 1173 ATAATTAGCGCTGCCATCACAGCATCTGTTCTTTTGGTGGATATGCTGATCATCTGGG 1232
Qy 342 AlaLysArgLeuLeuValTyrGluPheValProAsnAsnAsnLeuLeuLeuHisLeuHis 361
Db 1233 TCCAGAGGCTGCTTGTATTGTAATTTGTTTCCCAACACACATTTGGAATTTCCATTTG 1292
Qy 362 GlyGluGlyArgProThrMetGluTyrSerThrArgLeuLysLysLysLysLysSerAla 381
Db 1293 GGAAGAGGCGACCTTACATGATTTGGCCCAAGACTAAGAAATGCTTTTAGGATCTGCT 1352
Qy 382 LysGlyLeuSerTyrLeuHisGluAspCysAsnProLysLysLysLysLysLysLysLys 401
Db 1353 AAGGACTGGCGTATCTTCATGAAGATTTGTCATCTTAAGATCATCTCGTATCAAA 1412
Qy 402 AlaSerAsnIleLeuIleAspPheGluAlaLysValAlaAspPheGlyLeuAla 421
Db 1413 GCTCCCAACATCTTCTGGATTTTAAAGTTTGAAGCAAGGTTGAGATTTTGGTCTTGCA 1472
Qy 422 LysIleAlaSerAspThrAsnThrHisValSerThrArgValMetGlyThrPheGlyTyr 441
```


2213	TATGATGCTTTAGTGGATCCTCGCGCTGGGAAAGGACTTCTCAATCTAATAGATGGCAAGA	2272
523	MetValAlaCYbAlaAlaAlaCYbAlaArgHisSerAlaAlaArgArgArgProArgMetSer	542
2273	ATGATAGCCTGTGCAGCTGCATGTGTAGCCATTCTGCAGCTGTCGGCCACGATATAGT	2332
543	GlnIleValArgAlaLeuGluGluYAsnValSerLeuSerAspLeuAsnGluGlyMetArg	562
2333	CAGTCTGTTGGGGCTTTGGAGGGCAATGTCTCTTTGGAGGACCTTAATGAAGGTGTTCGG	2392
563	ProGlyGlnSerAsnValTyrSerSerTyrGlyGlySerThrAspTyrAspSerSerGln	582
2393	CCTGGCCATAGCCGCTCTTTGGGTCTATAC---AGCAGCTCCGATTACGATTCTGGCCAG	2449
583	TyrAsnGluAspMetLysLysPheArgLysMetAlaLeuGlyThrGlnGluTyrAsnAla	602
2450	TACACAGGAGNCATGAGAGAGTTCAAGAGATGGCATTCACAACAAC-----TATACC	2503
603	ThrGlyGluTyrSerAsnProThrSerAspTyrGlyLeuTyrProSerGlySerSerSer	622
2504	AGCAGCCAATACAGCGCGCCAAACCAGTGAATATGGACAGATACCGTCTGCATCAAGCAGC	2563
623	GluGlyGlnThrArgGluMetGluMetGlyLysIleLysArgThrGly-----	639
2564	GAGGGCCACCAGACGCAGGAGATGGAGTCCGGGTGCAATGAAGAAAGGTGGCTACAGTGGT	2623
640	-----GlnGlyTyrSerGlyProSer	646
2624	GGCTACAGCTCAGGATACAGCGGAGCCTCG	2653

RESULT 7

```

RES001
US-10-086-464-10
; Sequence 10, Application US/10086464
; Publication No. US20020199218A1
; GENERAL INFORMATION:
; APPLICANT: GORING, Daphne R. et al.
; TITLE OF INVENTION: PROLINE-RICH EXTENSIN-LIKE RECEPTOR KINASES
; FILE REFERENCE: P 25,762-A USA
; CURRENT APPLICATION NUMBER: US/10/086,464
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 10/069,304
; PRIOR FILING DATE: 2002-02-19
; PRIOR APPLICATION NUMBER: PCT/CA00/00966
; PRIOR FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: US 60/149,466
; PRIOR FILING DATE: 1999-08-19
; PRIOR APPLICATION NUMBER: US 60/159,122
; PRIOR FILING DATE: 1999-10-13
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 10
; LENGTH: 1902
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1) .. (1902)
US-10-086-464-10

```

Alignment Scores:		
Pred. No.:	5,75e-105	Length: 1302
Score:	1748.50	Matches: 367
Percent Similarity:	69.67%	Conservative: 81
Best Local Similarity:	57.08%	Mismatches: 141
Query Match:	50.64%	Indels: 55
DB:	14	Gaps: 16

US-10-086-464-2 (1-647) x US-10-086-464-10 (1-1902)

	QY	3 SerAlaProSerProGlyThrGlySerProSerProSerProSerAsnSerThrThrThr 22
	Dd	TCTGTCTCCTCCACAAACTCCACCCTTCTTCATCTCCACCGTTAATAACAATTCAACC 78

Qy		23	ThrProProAlaSerAlaProProProThrThrProSerSerProProProSer 42
Db		79	ACCTCTTCTCCG-----CGGGTCGGTCTCCTCTCTCTACACCTCTCAAGGA 129
Qy		43	ThrIleProThrSerProProSerSerArgSerThrProSerAlaProProProSer 62
Db		130	GACTCATCATCGCACCTCTGTATTCCATCTCCACCAGCTCCACNAGTCTCTAAC 189
Qy		63	ProProThrProSerThrProGlySerProProProLeuProGlnProSerProAla 82
Db		190	CCTCCTTAAT--TCCTTAATACTCTCTCTCCCTCCGCACACAGGCGGTGGAGGAA 246
Qy		83	ProThrThrProGlySerProProAlaProValThrProProThrArgAnProProPro 102
Db		247	AGAGGAAATGGAGGAAAACAATGGTGGAATGATACTCCACCGTCAACGGCTCTCTCCT 306
Qy		103	SerValProGlyProProSerAsnProSerArgGluGlyGlySerProArg-----Pro 120
Db		307	TCT-----CCTCCTCTTAGGAGTAATGGTAGAATGGTGGTAGCAGATCATCGCCA 357
Qy		121	ProSer-----SerProSerProProSerPro-----SerSerAsp 132
Db		358	CCAGGACACTGGAGGCTCTCGCTCAGACACCTCTCTTACTACGAGGAGGAGCAGTGGA 417
Qy		133	Gly-----LeuSerThrGlyValValValGlyLleAlaIleGly 145
Db		418	GGAGGTGGAGGTGGAGAAGTAATACGAAATACAGCGATCATAGTTGGTATTAGTCGGA 477
Qy		146	GlyValAlaLeuLeuValIleValThrLeuLeuCysLeuLeuCysIysIysIysArqArg 165
Db		478	GCTGGACTTTTGATGATCGTCTTATTATTGTGTGCTTTAGACGCCAAAAAGAGAGAAAA 537
Qy		166	ArgAspGluGluAspAlaTyTrValProProProProProGlyProIlysAlaGly 185
Db		538	-----GACTCTCTCTAC-----CCTGAACCCATGAAAGGA 567
Qy		186	GlyProTyrglyGlyngInGlnGlnTrpArgGlnGlnAsnAlaThrProProSerAsp 205
Db		568	AACCAATAT-----CAATACTA-TGGAAACAACAACAACAATGCTTTCACAGAA 617
Qy		206	HisValValThrSerLeuProProProProPolyeAlaProSerProPro-----221
Db		618	-----TTATCCGAATTTGGCACCTTAATTCACAAGGCCAAAAACAACATCTACTGTGG 671
Qy		222	-----ArgGln-ProProProProProPheMetSerSerGly-----237
Db		672	TTGGGGAGGGGTGGACCATCACCGCTCTCTCCCGGATGCCITACACGCGGAGGA 731
Qy		238	-GlySerAspTyrser-----AspArgProValLeuProProSerProGlyLeuVa 255
Db		732	TTCTTCCATGTACTCAGGCCCATCACGCCAGTTTACCTCTCTTCGCGCTGCTCTAGC 791
Qy		255	LleUglyPheSerIysSerThrPheThrTyrGluLeuAlaAlaArgAlaThrAsnGlyPh 275
Db		792	CCTCGSATTCAACAAGAGCACTTTTACTTACCAAGAGCTTGCCTGCTCAACAGAGGGGT 851
Qy		275	eSerGluAlaAsnLeuLeuGlyGlnGlyGlyPheGlyTyTrValHisIysGlyValLeuPr 295
Db		852	TACGGATGCTAACCTTTTGGACAGGAGGATTTGGGTATGTCTCAANAAGAGTCTTGCC 911
Qy		295	oSerGlyLysGluValAlaVallysGlnLeuLysValGlySerGlyGlnGlyGluArgGl 315
Db		912	TAGCGGNAAGAGTACAGTTAAGATTTTAAAGCGGGTAGCGACAAGAGAGAGAGGA 971
Qy		315	uPheGlnAlaGluValGluIlelleSerArgValHisHieArghisLeuValSerLeuVa 335
Db		972	GTTTCAAAGCTGAGTGCATATCATPATAGCCGTGTGCATCATCGTATCTGTCTTTGGT 1031
Qy		335	IgLyTyrcysteIleaGlyAlaLysArgLeuLeuValTyrgluPheValProAsnAsnAs 355
Db		1032	TGGATATTGCANACTGATGCAGAGATGTTGGTTTTATAGTTTGTCTTCTAACAAAC 1091
Qy		355	nLeuGluLeuHiIsleuHisGlyGluGlyArgProThrMetGluTrpSerThrArgLeuLy 375

```
Db 1092 TTTGCAATATCATCTTCATGGAAAAATCTTCCGGTAATGAGTCTCCACTAGTGTGCG 1151
Qy 375 sileAlaLeuGlySerAlaLysGlyLeuSerTyrLeuHisGluAspCysAsnProLysI1 395
Db 1152 TATCGCCTTAGTGCTGCGAAAGGACCTGCTTACCTTCACGAAGACTGCCATCTCGGAT 1211
Qy 395 eileHisArgAspIleLysAlaSerAsnIleLeuIleAspPheLysPheGluAlaLysVa 415
Db 1212 CATTCACCGGACATCAAGTCTGCAATATCTCTTGACTTCAACTTTGATGCTATGGT 1271
Qy 415 lAlaAspPheGlyLeuAlaLysIleAlaSerAspThrAsnThrHisValSerThrArgVa 435
Db 1272 GGCTGATTTTGGATTAGTAACTGATTAACATCTGATAACAACACTCATGTATCTACTCGTGT 1331
Qy 435 lMetGlyThrPheGlyTyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLy 455
Db 1332 GATGGGAACCTTCGGATATCTAGCTCCAGATATGCTTCAAGCGGTAAATTAACCGAGAA 1391
Qy 455 sSerAspValPheSerPheGlyValValLeuLeuGluLeuIleThrGlyArgArgProVa 475
Db 1392 ATCCGATGTTTCTTACGGAGTTATGTTATTCGAACTTATACTGGAAACACCGGT 1451
Qy 475 lAspAlaAsnValTyrValAspAspSerLeuValAspTrpAlaArgProLeuLeuAs 495
Db 1452 TGAT---AATAGCATCACCATGGAGACACCTTAGTAGATTGGGCTCGGCTCTTATGGC 1508
Qy 495 nArgAlaSerGluGlnGlyAspPheGlyLeuLeuAlaAspAlaLysMetAsnGlyTy 515
Db 1509 TCGGCGCTAGAGATGGAACCTTAAATGAGCTCGCAGATCGGAGGCTTGAAGCAACTA 1568
Qy 515 rAspArgGluGluMetAlaArgMetValAlaCysAlaAlaCysValArgHisSerAl 535
Db 1569 CAACCCGCAAGAATGCTCGAATGGTGACTTGTGCGCTGCTAGCATTCGTCTATTCGG 1628
Qy 535 aArgArgArgProArgMetSerGlnIleValArgAlaLeuGluGlyAsnValSerLeuSe 555
Db 1629 GCGTAAACGTCCTCAAGATAGCCAGATAGTAGAAGCGTTAGAAGGAGAGTGTCTTTAGA 1688
Qy 555 rAspLeuAsnGluGlyMetArgProGlyGlnSerAsnValTyrSerSerTyrGlyLys 575
Db 1689 TGCTTTAAACGAAGGTGTGAAGCCAGGACAGTAACGTTTACGGGTCATTGGGAGCAAG 1748
Qy 575 rThrAspTyrAspSerSerGlnTyrAsnGluAspMetLysLysPheArgLysMetAlaLe 595
Db 1749 CTCGATTTATAGTCAGACATCTTAACATTCGACATAGAGAAATTCAGACAGATAGTTT 1808
Qy 595 uGlyThrGlnGluTyr-----AsnAlaThrGlyGluTyrSerAsnProThrSerAs 612
Db 1809 GTCGAGCCAAGAAATCCCAAGTCAGTACTGTGAAGGAACATCTAGTAATGATCTAGAGA 1868
Qy 612 pTyrGly 614
Db 1869 TATGGGA 1875

RESULT 8
US-10-086-464-9
; Sequence 9, Application US/10086464
; Publication No. US20020199218A1
; GENERAL INFORMATION:
; APPLICANT: GORING, Daphne R. et al.
; TITLE OF INVENTION: PROLINE-RICH EXTENSIN-LIKE RECEPTOR KINASES
; FILE REFERENCE: P 25,762-A USA
; CURRENT APPLICATION NUMBER: US/10/086,464
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 10/069,304
; PRIOR FILING DATE: 2002-02-19
; PRIOR APPLICATION NUMBER: PCT/CA00/00966
; PRIOR FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: US 60/149,466
; PRIOR FILING DATE: 1999-08-19
; PRIOR APPLICATION NUMBER: US 60/159,122
; PRIOR FILING DATE: 1999-10-13
```

```
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 1939
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-086-464-9

Alignment Scores:
Pred. No.: 5,86e-105 Length: 1939
Score: 1748.50 Matches: 367
Percent Similarity: 69.67% Conservative: 81
Best Local Similarity: 57.08% Mismatches: 141
Query Match: 50.64% Indels: 55
DB: 14 Gaps: 16

US-10-086-464-2 (1-647) x US-10-086-464-9 (1-1939)
Qy 3 SerAlaProSerProGlyThrGlySerProSerProSerProSerAsnSerThrThr 22
Db 36 TCTGCTCTCCAAACAACCTCCACCTCTTCTCCATCTCCACCGTCTAATACCAATTCAC 95
Qy 23 ThrProProAlaSerAlaProProProThrThrProSerSerProProProSer 42
Db 96 ACCTCTTCTCCG-----CGGGCTCCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCA 146
Qy 43 ThrIleProThrSerProProSerSerArgSerThrProSerAlaProProProSer 62
Db 147 GACTCATCATCTCGCCACCTCTGATTCCACATCTCCACGATCTCCACAGCTCTCTTAAC 206
Qy 63 ProProThrProSerThrProGlySerProProLeuProGlnProSerProAla 82
Db 207 CCTCTTAAT---TCCTCTAATACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTT 263
Qy 83 ProThrThrProGlySerProProAlaProValThrProProThrArgAsnProPro 102
Db 264 AGAGGAAATGGAGAAACAATGTGGCAATGATATCTCCACGCTCCACGGCTCTCTCTCT 323
Qy 103 SerValProGlyProProSerAsnProSerArgGluGlyGlySerProArg-----Pro 120
Db 324 TCT-----CCTCTCTCTAGGAGTAATGGAGATAATGGTGTAGCAGATCATCGCCA 374
Qy 121 ProSer-----SerProSerProProSerPro-----SerSerAsp 132
Db 375 CCAGGACACTGGAGCTCTCGCTCAGACAACCTCTCTAGCGGAGGAGGAGGAGGAGG 434
Qy 133 Gly-----LeuSerThrGlyValValValValValValValValValValValVal 145
Db 435 GGAGGTGGAGGTGGAGAGAGTAATACGATACAGCGATCATAGTTGGTGTATTAGTCGGA 494
Qy 146 GlyValAlaLeuLeuValIleValThrLeuIleCysLeuLeuCysLysLysLysLysArg 165
Db 495 GCTGGACTTTTGTATGATGCTTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 554
Qy 166 ArgAspGluGluAspAlaTyrValProProProProProProProProProProPro 185
Db 555 -----GACTCTCTCTAC-----CCTGAACCCATGAAAGGA 584
Qy 186 GlyProTyrGlyGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 205
Db 585 AACCAATAT-----CAATACTA-TGGAACAACAACAACAACAACAACAACAACAACA 634
Qy 206 HisValValThrSerLeuProProProProProProProProProProProProPro 221
Db 635 -----TTATCCGAATGGCACCTTAATTCACAAGGCCAAACCAACAACAACAACA 688
Qy 222 -----ArgGln-ProProProProProProProProProProProPheMetSerSer 237
Db 689 TTGGGAGGGGGTGGACCATCATCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCA 748
Qy 238 -GlySerAspTyrSer-----AspArgProValLeuProProProSerProGlyLeuVa 255
Db 749 TTCTTCCATGTACTAGGCCCATCATCGCCCGAGTTTACCTCTCTCTCTCTCTCTCTCT 808
```

```
Qy 255 lLeuGlyPheSerIysSerThrPheThrTyrGluLeuAlaArgAlaThrAsnGlyPhe 275
Db 809 CCTCGGATTCAACAAGAGCACTTTTACTTACCAAGAGCTTCGGCTGCAACAGGAGGGTT 868

Qy 275 eSerGluAlaAsnLeuLeuGlyGlnGlyGlyPheGlyTyrValHisGlyValLeuPr 295
Db 869 TACGGATGCTAACCTTTTGGACAGGGAGGATTTGGGTATCTCCATTAAGGAGTCTTGCC 928

Qy 295 oSerGlyLysGluValAlaValLysGlnLeuLysValGlySerGlyGlnGlyGluArgG1 315
Db 929 TAGCGGGAAGAAGTAGAGATTAGAGTTTAAACCGGTAGCGGACAGAGAGAGGGA 988

Qy 315 uPheGlnAlaGluValGluLeuIleSerArgValHisHisArgHisLeuValSerLeuVa 335
Db 989 GTTTCAAGCTGAGGTGCGATATCAATTAGCGGTGTGCATCATCGGTATCTTGTCTTGGT 1048

Qy 335 lGlyTyrCysIleAlaGlyAlaLysArgLeuLeuValTyrGluPheValProAsnAsnAs 355
Db 1049 TGGATATTGCGATAGCTGATGACAGAGAGATGTTGGTTTATGAGTTGTTCCTAAACAAAC 1108

Qy 355 nLeuGluLeuHisLeuHisGlyGluGlyArgProThrMetGluTrpSerThrArgLeuLy 375
Db 1109 TTTGGATATCATCTTCATGGGAAAAATCTTCGGTAATGAGGTCTCCACTAGGTGCG 1168

Qy 375 sIleAlaLeuGlySerAlaLysGlyLeuSerTyrLeuHisGluAspCysAsnProLysIl 395
Db 1169 TATCGCCTTAGGTGCTGCGAAAGGACTCGCTTACCTTCACGAAGACTGCCATCCTCGGAT 1228

Qy 395 eIleHisArgAspIleLysAlaSerAsnIleLeuLeuAspPheLysPheGluAlaLysVa 415
Db 1229 CATTACCGCGCACATCAAGTCTGAAATATTCTCTTGGACTTCAACTTTGATGTATGGT 1288

Qy 415 lAlaAspPheGlyLeuAlaLysIleAlaSerAspThrAsnThrHisValSerThrArgVa 435
Db 1289 GGCTGATTTGGATTAGTAGTAATCAACATCTGATAACACACTCATGTATCTACTCGTGT 1348

Qy 435 lMetGlyThrPheGlyTyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLy 455
Db 1349 GATGGAACTTTCCGATATCTAGCTCCAGAAATATGCTTCAAGCGGTAAATTAACCGAGAA 1408

Qy 455 sSerAspValPheSerPheGlyValValLeuLeuGluLeuIleThrGlyArgArgProVa 475
Db 1409 ATCCGATGTTTCTCTTACGAGGATTATGTTATGGAACTTATAACTGGAACACACCGGT 1468

Qy 475 lAspAlaAsnValTyrValAspSerLeuValAspTrpAlaArgProLeuLeuAs 495
Db 1469 TGAT---AATAGCATCACCATGGACGACACCTTAGTAGTTGGCTCGGCTCTTATGGC 1525

Qy 495 nArgAlaSerGluGlnGlyAspPheGluGlyLeuAlaAspAlaLysMetAsnGlyTy 515
Db 1526 TCGCGCGCTAGAAAGTGAACCTTTAATGAGCTCGCAGATCGAGGCTTGAAGCAACTA 1585

Qy 515 rAspArgGluGluMetAlaArgMetValAlaCysAlaAlaCysValArgHisSerAl 535
Db 1586 CAACCCGCAAGAAATGGCTCGAATGGTGACTTGTGCCCTGCTAGCATTCGTCTATCGGG 1645

Qy 535 aArgArgArgProArgMetSerGlnIleValArgAlaLeuGluGlyAenValSerLeuSe 555
Db 1646 CGGTAAACGCTCCAAGATGACCGAGATAGTAAGAGCGTTAGAGGAGAAGTGTCTTAGA 1705

Qy 555 rAspLeuAsnGluGlyMetArgProGlyGlnSerAsnValTyrSerSerTyrGlyLysE 575
Db 1706 TGCTTTAAACGAAGGTGTGAAGCCAGGACACAGATAACGTTTACGGGTCAATGGGAGCAAG 1765

Qy 575 rThrAspTyrAspSerSerGlnTyrAsnGluAspMetLysPheArgLysMetAlaLe 595
Db 1766 CTCGGATTATAGTCAGACATCTTACAAATGCGAGACATGAAGAAATTCAGACAGATAGCTTT 1825

Qy 595 uGlyTyrGlnGluTyr-----AsnAlaThrGlyGluTyrSerAsnProThrSerAs 612
Db 1826 CTCGAGCCCAAGAAATTCGCCAGTCAGTGACTGTGGAAGGAACATCTAGTATGATCTTAGAGA 1885
```

```
Qy 612 pTyrGly 614
Db 1886 TATGGGA 1892
```

RESULT 9

```
US-10-425-114-9117
; Sequence 9117, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 9117
; LENGTH: 1424
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700833950_FLI
US-10-425-114-9117
```

```
Alignment Scores:
Pred. No.: 2,26e-104 Length: 1424
Score: 1737.50 Matches: 331
Percent Similarity: 88.73% Conservative: 47
Best Local Similarity: 77.70% Mismatches: 41
Query Match: 50.32% Indels: 7
Gaps: 4
```

US-10-086-464-2 (1-647) x US-10-425-114-9117 (1-1424)

```
Qy 226 ProProProProProPheMetSerSerGly-----GlySerAspTyrSerAsp 243
Db 12 CCCCCACCAACCGCTCTTTCATCAGCAGCAGTCTGGATCAACAAATTCAGGC 71

Qy 244 ArgProValLeuProProSerProGlyLeuValLeuGlyPheSerLysSerThrPhe 263
Db 72 GGTGAATTTCTCTCTCTCTCTCCAGGAATTTCAATTGGGGTTCTCTAAGAGCACATTC 131

Qy 264 ThrTyrGluGluLeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuGlyGln 283
Db 132 ACGTATGAGGAGTTGGCAGCGCAACTGATGGCTTCTGTATGCCAACCTCTTGGACAA 191

Qy 284 GlyGlyPheGlyTyrValHisLysGlyValLeuProSerGlyLysGluValAlaValLys 303
Db 192 GGAGGATTTGATATGTGCACAGAGGAATTTCTCCCAACGGCAAGGAGTGGCAGTGAAG 251

Qy 304 GlnLeuLysValGlySerGlyGlnGlyGluArgGluPheGlnAlaGluValGluIleIle 323
Db 252 CAATTGAAGGCTGGAAGCGGCAAGCGTGAATTTCCAAGCTGAAGTTGAGATAAAT 311

Qy 324 SerArgValHisHisArgHisLeuValSerLeuValGlyTyrCysIleAlaGlyValLys 343
Db 312 AGCGGTGTCCATCAAGACATCTTTGTTTCTTTGGTGTGATCTGATCATCTGCTGGTCCAG 371

Qy 344 ArgLeuLeuValTyrGluPheValProAsnAsnAsnLeuGluLeuHisLeuHisGlyGlu 363
Db 372 AGGCTGCTGTTTATGAATTTGTTTCCCAACACACATTTGGAATTTCCATTTGATGAAAA 431

Qy 364 GlyArgProThrMetGluTrpSerThrArgLeuLysIleAlaLeuGlySerAlaLysGly 383
Db 432 GGGAGACTACCATGGATTCGCCACAGACTAAGAAATTCCTTTAGGATCTCTGCTAAGGGA 491

Qy 384 LeuSerTyrLeuHisGluAspCysAsnProLysIleIleHisArgAspIleLysAlaSer 403
```

Db 492 CTGGCGTATCTTCATGAAGATTGTTCATCTCAAGATCATCCATCGTGATATCAAAATCTGCC 551
Qy 404 AenileLeuIleAaspPheGluAlaLysValAlaAaspPheGlyLeuAlaLysIle 423
Db 552 AACATCCTCTGGATTGATTTGAAGTGAACAAAGGTTGCAGATTTTGGTCTTGCAGAGTTT 611
Qy 424 AlaSerAaspThrAsnThrHisValSerThrArgValMetGlyThrPheGlyTyrLeuAla 443
Db 612 TCTTCTGATGCAATACCCATGTTTCTACTCGAGTGATGGGACATTTTGGGTATTTGGCT 671
Qy 444 ProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAaspValPheSerPheGlyVal 463
Db 672 CCAGAATATGCTTCTAGTGGAAATCTACAGCAAAATCAGATGTTTCTCTATAGGAATC 731
Qy 464 ValLeuLeuGluLeuIleThrGlyArgArgProValAaspAlaAsnAsnValTyrValAasp 483
Db 732 ATGCTCCTTGAGTTAATACCGGACGCGGCGGAGTTGATTAANAATCAAACTTACATGGAG 791
Qy 484 AspSerLeuValAaspTrrpAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAaspPhe 503
Db 792 GATAGTTTGGTAGATTGGCTAGGCTTTGCTCACACGAGCTTTTGGAAAGAGGATGATTTT 851
Qy 504 GluGlyLeuAlaAaspAlaLysMetAsnGlyTyrAaspArgGluGluMetAlaArgMet 523
Db 852 GATTCATTATTGACCCCAAGGCTCCAGAAATGACTATGATCTCTCATGAGATGGCACGAATG 911
Qy 524 ValAlaCysAlaAlaCysValArgHisSerAlaArgArgArgProArgMetSerGln 543
Db 912 GTGGCTTCTGCGGCTTGATAGCTCATTCGGCAAGCGTTCGACCAAGGATGAGCCAG 971
Qy 544 IleValArgAlaLeuGluGlyAsnValSerLeuSerAaspLeuAsnGluGlyMetArgPro 563
Db 972 GTTGTCCGCGCTCTCGAAGGAGATGCTCTCTAGCAGATCTGAATGAAGGAATAAGACCT 1031
Qy 564 GlyGlnSerAsnValTyrSerTyrGlyGlySerThrAaspTyrAaspSerSerGlnTyr 583
Db 1032 GGACACACACTATGATATGTTCTCAT---GAAAGCTCAGATTTATGACACTGCACAGTAC 1088
Qy 584 AsnGluAaspMetLysLysPheArgLysMetAlaLeuGlyThrGlnGluTyrAsnAlaThr 603
Db 1089 AAGGAAGACATGAANAAGTTCAGGANAATGGATTGGGAATTCAGGAGTATGGTGCAAGC 1148
Qy 604 GlyGluTyrSerAsnProThrSerAaspTyrGlyLeuTyrProSerGlySerSerGlu 623
Db 1149 AGTGAGTACAGTCCGCTACAGTGAGTATGTTTAAACCCATCAGGCTCAAGTAGTAA 1208
Qy 624 Gly-----GlnThrArgGluMetGluMetGlyLysIleLysArgThrGlyGln 640
Db 1209 GCACAGAGCCGCAAAACCAAGGGAATGGAATGAGAAAGATGAAG---AACAAATCAA 1265
Qy 641 GlyTyrSerGlyProSer 646
Db 1266 GGTTCAGTGAAGTTCT 1283

RESULT 10

US-10-086-464-13
; Sequence 13, Application US/10086464
; Publication No. US20020199218A1
; GENERAL INFORMATION:
; APPLICANT: GORING, Daphne R. et al.
; TITLE OF INVENTION: PROLINE-RICH EXTENSIN-LIKE RECEPTOR KINASES
; FILE REFERENCE: P 25,762-A USA
; CURRENT APPLICATION NUMBER: US/10/086,464
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 10/069,304
; PRIOR FILING DATE: 2002-02-19
; PRIOR APPLICATION NUMBER: PCT/CA00/00966
; PRIOR FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: US 60/149,466
; PRIOR FILING DATE: 1999-08-19
; PRIOR APPLICATION NUMBER: US 60/159,122
; PRIOR FILING DATE: 1999-10-13
; NUMBER OF SEQ ID NOS: 27

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 2025
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2025)
US-10-086-464-13

Alignment Scores:
Pred. No.: 6,08e-101 Length: 2025
Score: 1687.00 Matches: 368
Percent Similarity: 63.99% Conservative: 142
Best Local Similarity: 51.76% Mismatches: 87
Query Match: 48.86% Indels: 114
DB: 14 Gaps: 25

US-10-086-464-2 (1-647) x US-10-086-464-13 (1-2025)

Qy 1 MetSerSerAlaProSerProGlyThrGlySerProProSerProProSerAsnSerThr 20
Db 16 GTGATTCATCTCTGCCCCCTGAAACC---TCAANTGGGACACCCCGTCANACGGAACA 72
Qy 21 ThrThrThrProProAlaSerAlaProProThrThrProSerSerProProPro 40
Db 73 TCG-----CGTCTAATGAGTCATCGCGGCCAACACCCACCTTCTTACACCAACCA 123
Qy 41 ProSerThrThrProThrProProSerSerSerArgSerThrProSerAlaProPro 60
Db 124 TCATCA-----ATATCTGCTCTCCGCGCAGATATCTCCGCTTCTTTTTCACCGCG 174
Qy 61 ProSerProThr-----ProSerThrProGlySerProProProLeu 75
Db 175 CTGCACCAACGCAAGAAACGTCACCTCTACATCTCCGTCCTCTATCCGCCCTGTT 234
Qy 76 -----ProGln-----ProSerProAlaPro-----Thr 84
Db 235 GTAGCTAATCCGTCACCGCAGACTCCACAGAAATCTTCTCCACCTGCACCTGAAGGCTCA 294
Qy 85 ThrProGlySerProAla---ProValThrPro-----ProThrArgAsn 99
Db 295 ACTCCTTAACGCGCACCTGCACCAACCAACCAACCGTCGAACCAATCACCAGGAAGACCA 354
Qy 100 ProProPro-----Pro-----Pro-----Pro----- 102
Db 355 ACTCCTCTCTCTCTCTGTCGCAATGATGACCGAACAAGAACCAATGCGGGAACAACAC 414
Qy 103 -----SerValProGlyProProSerAsnProSerArg-----GluGlyGly 116
Db 415 AGAGACGCTCCACACCATCACCCCGTCGTCAGGGAACAGAACTTCCGGTGACGGTGGC 474
Qy 117 SerProArgProProSerProProSerProProSerProSerSerAspGlyLeuSerThr 136
Db 475 TCACCTTTCACACCTCGGTGATAAGCCCTCTCAGAAATAGTGGAGATTACAGACTCATCA 534
Qy 137 GlyValValGlyIleAlaIleGlyGlyValAlaLeuValIleValThrLeuIle 156
Db 535 TCG-----GGGCTT-----TTGCTTCTTACTTGCAGTGTGTATT 567
Qy 157 CysLeuLeuCysLysLysLysArgArgArgAspGluAspAla-----Tyr 172
Db 568 TGCATCTGTTGCAACAGAGAGAGAGAGAAATCTCTCAGGTCAACCACTCACTAC 627
Qy 173 TyrValProProProProProGlyProLysAla---GlyGlyProTyrGlyGlyGln 191
Db 628 TACAATAACAATCCTTATGGAGGAGCACCTCAGGTAATGCTGTTATTATACAGGGA--- 684
Qy 192 GlnGlnGlnTrpArgGlnGlnAsnAlaThrProProSerAspHisValValThrSerLeu 211
Db 685 -----ACACCTCAAGATCATGTGGTG---AATATG 711
Qy 212 ProProProProLysAlaProSerProProArgGlnProProProProProProPro 231

Db 712 GCTGTCAGGAGGTGGGAATTTGGGTCACAGCAACCTGTGTCTGGTCCT----- 762
Qy 232 PheMetSerSerGlyGlySerAspTyrSerAspArgProValLeuProProSer 251
Db 763 -----CACAGTGATGCTTCCAACTTAACCGGTGGAATGCTATACCGTCACCTCAA 813
Qy 252 ProGlyLeuValLeuGlyPheSerLysSerThrPheThrTyrGluGluLeuAlaArgLa 271
Db 814 ---GCTGCAACTCTTGGTCACAAACCAAGCACTTTTCATACGATGGAAGTGTCCATTGCA 870
Qy 272 ThrAenGlyPheSerGlnAenLeuGlyGlnGlyGlyPheGlyTyrValHisLys 291
Db 871 ACAGNAGGTTTCGCTCAGTCAGTCAAAATTTGTAGGCAAGGAGGATTTGGGTATGTTTCATA 930
Qy 292 GlyValLeuProSerGlyGlyGluValAlaValLysGlnLeuLysValGlySerGlyGln 311
Db 931 GGAGTTCCTGCTAGTGGCAAGAGTTCAGTCAGTGAAGAGTCTTAAACCTTGGGAAGTGGACA 990
Qy 312 GlyGluArgGluPheGlnAlaGluValGluLleIleSerArgValHisArgHisLeu 331
Db 991 GGGGAACCGAGTTTCAAGCAGAGGTTGATATCATTTAGCCGTGTCATCATCGTCATCTC 1050
Qy 332 ValSerLeuValGlyTyrCysIleAlaGlyAlaLysArgLeuLeuValTyrGluPheVal 351
Db 1051 GTTTCCTCTGTGGATATGATCTCTCGTGGTCAAGACATTTGGTTATGAGTTTATA 1110
Qy 352 ProAenAenLeuGluLeuHisLeuHisGlyGlyGluArgProThrMetGluTrpSer 371
Db 1111 CCTAACACACTCTTGAATTTTCATCTTCATGGAAGGTCGTCGGTTCGTTTGGATTCGCT 1170
Qy 372 ThrArgLeuLysIleAlaLeuGlySerAlaLysGlyLeuSerTyrLeuHisGluAspCys 391
Db 1171 ACAAGATGAAGATTGCATTCGATTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTC 1230
Qy 392 -----AenProLysIleIleHisArgAspIle 400
Db 1231 AAGAAAACTTTATCTCATATATTTGTCATCAGTCACCTCGCATTCATTCACAGAGATATC 1290
Qy 401 LysAlaSerAenIleLeuLeuAspPheLysPheGluAlaLysValAlaAspPheGlyLeu 420
Db 1291 AAAGCTGCAACATCTTCTTCTGATTCAGTTTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGT 1350
Qy 421 AlaLysIleAlaSerAspThrAsnThrHisValSerThrArgValMetGlyThrPheGly 440
Db 1351 GCTAAGCTATCTCAAGACAACATATATCTCATGTCCTCCACTCGCTCATGGCACTTTTGG 1410
Qy 441 TyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPheSer 460
Db 1411 TACTTAGCTCCAGAGTATGCATCAAGCGGAAGTTATCCGACAAATCTGATGTTTCTCA 1470
Qy 461 PheGlyValValLeuLeuGluLeuIleThrGlyArgArgProValAspAlaAenVal 480
Db 1471 TTTGGAGTAATGCTTCTTGGAGCTATAACCGGAAGACCTCTCTCGATCTAACTGGA--- 1527
Qy 481 TyrValAspAspSerLeuValAspTrpAlaArgProLeuLeuAenArgAlaSerGluGln 500
Db 1528 GAAATGGGAAGATAGCTTGGTATGGCAAGGCGCTTTGTGTGTTGAAAGCAGCTCAAGAT 1587
Qy 501 GlyAspPheGluGlyLeuAlaAspAlaLysMetAenGlyTyrAspArgGluGluMet 520
Db 1588 GGAGATTACACCAATTTGGTGATCCAGTCTAGAGCTAAACTACAGTCATCAAGAGATG 1647
Qy 521 AlaArgMetValAlaCysAlaAlaCysValArgHisSerAlaArgArgProArg 540
Db 1648 GTTCAATGGCTTCTTGTGACGTCAGCAATCAGACATTCAGCAAGAAGACGGCTAAG 1707
Qy 541 MetSerGln-----IleValArgAlaLeuGlu 549
Db 1708 ATGACCCAGGTTCAAAAACTCATACCTTGTGTTCTTATTATTATTGTCAGGACTAGAA 1767
Qy 550 GlyAenValSerLeuSerAspLeuAenGluGlyMetArgProGlyGlnSerAenValTyr 569

Db 1768 GGAGATATGTCAATGGATGATCTAAGTGAGGGAACAAGACCAGGACAAACGACGTACTTG 1827
Qy 570 SerSerTyrGlyGlySerThrAspTyrAspSerSerGlnTyrAenGluAenMetLysLys 589
Db 1828 AGCCCGGGAGCGGTGAGCTCAGAGTATGACCAAGCTCGTACACGGCAGACATGAAAAAG 1887
Qy 590 PheArgLysMetAlaLeuGlyThrGlnGluTyrAenAlaThrGlyGluTyrSerAenPro 609
Db 1888 TTCAAGAACTGGCGTTAGAGATAAAGATATCAAGCAGT---GAATATGGTGA--- 1941
Qy 610 ThrSerAspTyrGlyLeuTyrProSerGlySerSerGluGlyGlnThrThrArgGlu 629
Db 1942 ACAAGTCAGTATGCTTAAACCCCTTCTCTTCAAGTAGTGAA-----GAA 1986
Qy 630 MetGluMetGlyLysIleLysArgThrGlyGln 640
Db 1987 ATGAATAGAGCTCAATGAACGCAATCCTCAG 2019
RESULT 11
US-10-086-464-12
; Sequence 12, Application US/10086464
; Publication No. US20020199218A1
; GENERAL INFORMATION:
; APPLICANT: GORING, Daphne R. et al.
; TITLE OF INVENTION: PROLINE-RICH EXTENSIN-LIKE RECEPTOR KINASES
; FILE REFERENCE: P 25 762-A USA
; CURRENT APPLICATION NUMBER: US/10/086,464
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 10/069,304
; PRIOR FILING DATE: 2002-02-19
; PRIOR APPLICATION NUMBER: PCT/CA00/00966
; PRIOR FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: US 60/149,466
; PRIOR FILING DATE: 1999-08-19
; PRIOR APPLICATION NUMBER: US 60/159,122
; PRIOR FILING DATE: 1999-10-13
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 2104
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-086-464-12
Alignment Scores:
Pred. No.: 6,31e-101 Length: 2104
Score: 1687.00 Matches: 368
Percent Similarity: 63.99% Conservative: 87
Best Local Similarity: 51.76% Mismatches: 142
Query Match: 48.86% Indels: 114
Gaps: 25
US-10-086-464-2 (1-647) x US-10-086-464-12 (1-2104)
Qy 1 MetSerSerAlaProSerProGlyThrGlySerProSerProSerProSerAenSerThr 20
Db 52 GTGGATTCACTCTCTGCTCCCTGAAACC---TCAATGGGACACCAACCGTCAACGGAACA 108
Qy 21 ThrThrThrProProAlaSerAlaProProThrThrProSerSerProProPro 40
Db 109 TCG-----CCGTCTAATAGTCATCGCGCGCAACACCAACCTTCTTCAACCAACCA 159
Qy 41 ProSerThrIleProThrSerProSerSerArgSerThrProSerAlaProPro 60
Db 160 TCATCA-----ATATCTGCTCTCCGCGCAGATATCTCCGTTCTTTTTCACCGCG 210
Qy 61 ProSerProProThr-----ProSerThrProGlySerProProProLeu 75
Db 211 CTGCACCAACCAACGCAAGAAACGTCACCTCTACATCTCCGTCCTCATCGCGCTGT 270
Qy 76 -----ProGln-----ProSerProAlaPro-----Thr 84
Db 271 GTAGCTAATCCGTCAACCGCAGACTCCAGAGAATCTTCTCCACCTGCACCTGAAGGCTCA 330

Db 1949 AAGAACTACGATAGAAATGTTTCGGATGATCGAGGTCGCCAGCCTGTGTACGC 2008
Qy 533 HisSerAlaArgArgProArgMetSerGlnIleValArgAlaLeuGluGlyAsnVal 552
Db 2009 CACTCATCGTGAAGAGACCGCATGAGTACGGTGGTGAGAGCTTTAGATTCCTTGGAT 2068
Qy 553 SerLeuSerAspLeuAsnGluGlyMetArgProGlyGlnSerAsnValTyrSerSerTyr 572
Db 2069 GAGTTTACGGATCTCAATAACGGAATGAACCGGCACAGAGTTCGGTG----- 2116
Qy 573 GlyGlySerThrAspTyrAspSerSerGlnTyrAsnGluAspMetLysLysPheArgLys 592
Db 2117 -----TTTGATTCGGCGCAATCTGCACAAATCAGAATGTTAGGAGG 2161
Qy 593 MetAlaLeuGlyThrGlnGluTyrAsnAlaThrGlyGluTyrSerAsnProThrSerAsp 612
Db 2162 ATGGCTTTTGGAGGCAAGAT-----AGTCCCGGTTTCTCAATGAGTCTCAGAGTAGC 2215
Qy 613 Tyr-GlyLeuTyrProSerGlySerSerSerGluGlyGlnThrThrArgGluMetGluMe 632
Db 2216 TGGAGGAGTA-----GAGATCAGCAGCCCAACAACACTGTCTTC 2251
Qy 632 tGlyLysIleLysArgThrGlyGlnGly 641
Db 2252 TCCCAAAATAAAA-----CTGGGCATGGGA 2276

RESULT 13

US-10-425-114-7339
; Sequence 7339, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 7339
; LENGTH: 2655
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700652889_FLI
US-10-425-114-7339

Alignment Scores:
Pred. No.: 3 3e-91 Length: 2655
Score: 1539.00 Matches: 341
Percent Similarity: 55.19% Conservative: 95
Best Local Similarity: 43.16% Mismatches: 141
Query Match: 44.57% Indels: 213
DB: 13 Gaps: 19

US-10-086-464-2 (1-647) x US-10-425-114-7339 (1-2655)

Qy 2 SerSerAlaProSerProGlyThrGlySerProProSerProProSerAsnSer----- 19
Db 98 TCATCTCTCCACA-----AACACTCCACCACATCACTCCAGTCTCTCAGCCT 154
Qy 20 -----ThrThrThrThrProProProAlaSerAlaPro-----ProProThrThr 34
Db 155 AATCAACACAAACAAACAAACAAACCAATTCCTCTCTCATCTCTCTCTGAC 214
Qy 35 ProSerSerProPro-----ProProSerThrIleProThrSerProProProSerSer 52
Db 215 CTTCTGTCCACCT 274

Qy 53 ArgSerThrPro-----SerAlaProProProSerProProProThrProSerThrPro 69
Db 275 TCACCTGTCTCTCTCCACCATCATCTACACCAACCACTTTGCCCTCCATCACCACCT 334
Qy 70 GlySerProProProLeuPro-----GlnPro 78
Db 335 GATTCTCCACCACTTGCACCTGTCATCTCCAAACCCACCACTTACAAACATCTCCCCCC 394
Qy 79 SerProProAlaProThrProGlySerProProAlaProValThrProProThrArg 98
Db 395 TCACCAACAGTTACAACATCTCCCTCCCTCACCACAGTACCACATCCCTCTCTCCAGCT 454
Qy 99 AsnProProProSerValPro-----GlyProProSerAsnProSerArgGluGly 116
Db 455 GAAACTCCACCTCTCTCTCCAAATTTGCTCTCCACCACTCTCCACCACTCTCCAGCT 505
Qy 117 SerProArgPro-----ProSerSerPro 124
Db 506 TCCCTCTCTCTCAATCACCTCCCAACAATTCACCTCTCTCTCTCGACCGATTTCCGCA 565
Qy 125 SerProProSerProSer----- 130
Db 566 TCTCTCTCTCTCGGCCAAATTTTCCAAAGGCCCAACCACTAGAACTCTCTCCAGAGAAA 625
Qy 130 ----- 130
Db 626 GAAATCCACCAAAAACCTACTCTCTTCACATGCATCTCTCCATCAGTTTCTGAAACTCCT 685
Qy 130 ----- 130
Db 686 CCTAAACCTCTCTCTCTGTATGTTCCCTCCCTCCACATTCCTTCAACTCTCTCTCTCA 745
Qy 130 ----- 130
Db 746 GACCTTTCAGATCTTCACTCTCCAGCTTCTTTGCTGATCTCTCCCACTAATAAACAAGTG 805
Qy 130 ----- 130
Db 806 GTGGGGGTCCGAAGGTGTCGTACCTCTCTTCCAACTGAGAAACCACTCTCTCTCTCTCT 865
Qy 131 -----SerAspGlyLeu 134
Db 866 ACTAATGATGGTACTTAACAGTATGTCTTCAAAACACACACCTTCACTCTCTCTCTCTCT 925
Qy 135 SerThrGly-----ValValValGlyIleAlaIleGlyGlyValAlaLeuValIle 152
Db 926 AGCACTGGAGATCTGTGGCTATTGGAAATTTGTTGTTTATTTGTTCTCTCTCTCTCTCT 985
Qy 153 ValThrLeuIleCysLeuLeuCysLysLysLysArgArgArgAspGluGluAspAlaTyr 172
Db 986 GTTATGGCTGTGTGTTTGCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1045
Qy 173 TyrValProProProProProProGlyProGlyProGlyProGlyProGlyProGlyGln 192
Db 1046 TATGCTCTCTCTCTCTCA----- 1063
Qy 193 GlnGlnTrpArgGlnGlnAsnAlaThrProProSerAspHisValValThrSerLeuPro 212
Db 1064 -----TTTACCTCTATCCCAATTCAGGTACCTTATTC 1096
Qy 213 ProProProLysAlaProSerProProArgGlnProProProProProProProPhe 232
Db 1097 TTGAGGCGGAGTCTCGGCC-----AACTTT 1123
Qy 233 MetSerSerSerGlyGlySerAspTyrSerAspArgProValLeuProProProSerPro 252
Db 1124 TTAGTAGTGGCTCTGGTAGTATTT-----GTATATTCTCCATCAGAGCT 1171
Qy 253 GlyLeuValLeuGlyPheSerLysSerThrPheThrTyrGluGluLeuAlaArgAlaThr 272
Db 1172 GGTGGTGA---AGTAGTTCAAGATCATGTTTACATATGAAGAACTATTATCAAGTACA 1298
Qy 273 AsnGlyPheSerGluAlaAsnLeuGlyGlnGlyPheGlyTyrValHisLysGly 292


```
Qy 146 GlyValAlaLeuLeuValIleValThrLeuIleCysLeuLeuCysLysLysLysArg 165
Db 530 TTCGCTCTGCTTCTTCCATCGTGGCGTGTGTGTCGCGCAAGAAGAAGAAAGT 589
Qy 166 ArgAspGluGluAspAlaTyrTyrValProProProProProGlyPro----- 182
Db 590 -----CCGCCCCCAATGAACATGCCCTTCTACACC 619
Qy 183 ---LysAlaGlyGly-----ProTyrGlyGlyGlnGlnGlnTrpArgGln 198
Db 620 GACGAGAAGCAATGTATTACCCCNACGCTGTCTGCCGCTATGTGCACCAATAT 679
Qy 199 AsnAlaThrProProSerAspHisValValThrSerLeuProProProProLysAlaPro 218
Db 680 GCGAGCAAC-----GGCAGCATCCCTCCGCCGG-ATGCGACCA 717
Qy 219 SerProProArgGlnProPro-ProProProProProPheMetSerSerGlyGly 238
Db 718 TCACGGCGCGGGAACCGCTGTGCCCATCGCTCGGCTCGATGGCGCGCCCTTGAGCGG 777
Qy 238 ySerAspTyrSerAsp-----ArgProValLeuProProProSerProGlyLeuVa 255
Db 778 CGAGATGACTCGTCCGCCGCCCGCCCGGCTTCCCTTCCCGCCACTTCCGCCAAGTGGC 837
Qy 255 lLeuGlyPheSerLysSerThrPheThrTyrGluGluLeuAlaAlaArgAlaThrAsnGlyPh 275
Db 838 GCTGGGGTTCTCCAAGAGCTCTTCTCGTACGAGGAGCTAGCGCGAGCGTCTGGGCTT 897
Qy 275 eSerGluAlaAsnLeuLeuGlyGlnGlyGlyPheGlyTyrValHisLysGlyValLeuPr 295
Db 898 CTCGTCGGCGAACCTGCTGGCGGCGGCGGCTTCCGGTACGTGTACAAAGGCGTGTGCG 957
Qy 295 o---SerGlyLysGluValAlaValLysGlnLeuLysValGlySerGlyGlnGlyGluAr 314
Db 958 CGGCGAGCGGGAAGAGTGGCGGTGAAGCACTCAAGTCCGGGAGCGCGGAGGGAGCG 1017
Qy 314 gGluPheGlnAlaGluValGluIleIleSerArgValHisHisLysLeuValSerLe 334
Db 1018 CGAGTTCAGCGGAGGTGGAGATCATCAGCGCGGTGCACCCACCTGCTACCTGGTGTGCT 1077
Qy 334 uValGlyTyrCysIleAlaGlyAlaLysArgLeuLeuValTyrGluPheValProAsnAs 354
Db 1078 CGTGGGCTACTGTCATCGCGGCAACACGCGCATGCTGCTACGAGTTCGTGGGCAACA 1137
Qy 354 nAsnLeuGluLeuHisLeuHisGlyGlyGlyArgProThrMetGluTrpSerThrArgLe 374
Db 1138 CACGCTGGGACCATCTGTACGCCAGGAGCGGCTTCTATGAGCTGGACACCCGCT 1197
Qy 374 uLysIleAlaLeuGlySerAlaLysGlyLeuSerTyrLeuHisGluAspCysAsnProLy 394
Db 1198 GAAGATCGGCTCGGCTCCGCGCAAGGGCTCGCTACTCTCCAGCGAGCTGCCATCTCG 1257
Qy 394 sIleIleHisArgAspIleLysAlaSerAsnIleLeuLeuAspPheLysPheGluAlaLy 414
Db 1258 GATCATCCCGCGACATCAAGGCGCGCAACATCTGCTGTGCACCAACACTTCGAGGCCAT 1317
Qy 414 sValAlaAspPheGlyLeuAlaLysIleAlaSerAspThrAsnThrHisValSerThrAr 434
Db 1318 GGTGGCGGACTTCGGGCTGGCAAGCTCACCGCGGACACCAACACGCGCTCTCCAGCG 1377
Qy 434 gValMetGlyThrPheGlyTyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGl 454
Db 1378 CGTCATGGCACCTTCGCTACTCTGGCGCGGAGTACGGCTCCAGCGGCAAGCTACCGGA 1437
Qy 454 uLysSerAspValPheSerPheGlyValValLeuLeuLeuLeuThrGlyArgArgPr 474
Db 1438 CAGTCGGGACGCTCTCTTCGCGGCTCATGCTCTGGAGCTACTCACGGGCGCGCGCC 1497
Qy 474 oValAspAlaAsnAsnValTyrValAspAspSerLeuValAspTrpAlaArgProLeuLe 494
Db 1498 CATGATACCAACCAAC---TACATGGAGGACAGCTCTGTGAGCTGGCGCGCGCTGCT 1554
Qy 494 uAsnArg-----AlaSerGluGlnGlyAspPheGluGlyLeuAlaAspAlaLysMetAs 512
```

```
Db 1555 GAGCGCGCGCTGGCGGGGAGACGGC---TTGCGCGAGCTCGTCGACCCCGCTCTGGG 1611
Qy 512 nAsnGlyTyrAspArgGluMetAlaArgMetValAlaCysAlaAlaCysValar 532
Db 1612 AGCGAGTACTCGCTCGTGGAGTGGAGCTTGGCGGCTTGGCGCGCGCCGCGCACACCG 1671
Qy 532 gHisSerAlaArgArgProArgMetSerGlnIleValArgAlaLeuGluGlyAsnVa 552
Db 1672 CCACTCGGCCAAGCGCGCCCAAGATGAGCCAGATCGTGGCGCGCTGGAGGGGACGC 1731
Qy 552 lSerLeuSerAspLeuAsn---GluGlyMetArgProGlyGlnSerAsnValTyrSerSe 571
Db 1732 ATCGCTGGAGATCTGCACGAGGACGGCGGCAAGCCGCGGCGAGCGTGTCTTCTTC- 1789
Qy 571 rTyrGlyGlySerThrAspTyrAspSerSerGlnTyrAsnGluAspMetLysLysPheAr 591
Db 1790 ---GGCGGTGGTCCGAC-----AACATCTCCCGCTTAAG 1821
Qy 591 gLysMetAlaLeuGlyThrGlnGluTyrAsnAlaThrGlyGluTyrSerAsnProThrSe 611
Db 1822 GCAGCTCGCTTCGACAGCGCGGACTACTCG-----GACTACAGCAGCGACTCCTC 1872
Qy 611 rAspTyrGly 614
Db 1873 CACCGTGGT 1882
```

RESULT 15

```
US-10-425-114-6300
; Sequence 6300, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCES: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 6300
; LENGTH: 1500
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700560157_FLI
US-10-425-114-6300
```

```
Alignment Scores:
Pred. No.: 8,16e-85 Length: 1500
Score: 1437.00 Matches: 284
Percent Similarity: 74.27% Conservative: 45
Best Local Similarity: 64.11% Mismatches: 80
Query Match: 41.62% Indels: 34
DB: 13 Gaps: 7
```

US-10-086-464-2 (1-647) x US-10-425-114-6300 (1-1500)

```
Qy 185 GlyGlyProTyrGlyGlyGlnGlnGlnTrpArgGlnGlnAsnAlaThrProProSer 204
Db 11 GGAGGTCCACATGGG----- 25
Qy 205 AspHisValValThrSerLeuProProProProLysAlaProSerProProArgGlnPro 224
Db 26 GACCATGTTGTG---AGGATGCACAAATGGATGGTCTCTGGTGGCGGAGGTGGGGT 82
Qy 225 ProProProProProProPheMetSerSerGly-----GlySerAspTyrSer 242
Db 83 GCACCAACCACTCTCTCCAATGATGATGAGTAGTCTGAGTTTAGCTCCAATTACTCG 142
```

```
Qy 243 AspArgProVal----LeuProProProSerProGlyLeuValLeuGlyPheSerLysSer 261
Db 143 ACGGACCGCGCCGCTTGCCTCCTCCCTCACCACAACTTGTCTGGGCTCAAAAGGGGCG 202
Qy 262 ThrPheThrTyrGluGluLeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeu 281
Db 203 ACCCTTCACTTACGAGGAATTAGCAGCTGCCACCAACGGATTCAAATGATGCAAAATTTGATA 262
Qy 282 GlyGlnGlyGlyPheGlyTyrValHisLysGlyValLeuProSerGlyLysGluValAla 301
Db 263 GGACAGGTGGATTCGGCTATGTCATAAGGGTGTGTGCTAGTGGAAAGAAAGTGGCA 322
Qy 302 ValLysGlnLeuLysValGlySerGlyGlnGlyGluArgGluPheGlnAlaGluValGlu 321
Db 323 GTTAAGAGCTTAAAGCAGGTAGTGGCCAGAGAGAGAGAGAGTTCCAAGCTGAGATTGAC 382
Qy 322 IleIleSerArgValHisHisArgHisLeuValSerLeuValGlyTyrCysIleAlaGly 341
Db 383 ATCATTAGCCGGTCCATCATCGCCATCTCTGTGCACCTGTGTGGATCTCAATTTCTGGT 442
Qy 342 AlAlysArgLeuLeuValTyrGluPheValProAsnAsnLeuLeuGluLeuHisLeuHis 361
Db 443 GGCCAGAGGATGTTGCTATGAATTTATTCCCAATAACACATTGGAAATATCACCTCCAT 502
Qy 362 GlyGluGlyArgProThrMetGluTrpSerThrArgLeuLysIleAlaLeuGlySerAla 381
Db 503 GGAAGGGTAGGCCCTACCATGGCTGGCCAACTAGAAATCGGATTCGAATAGGATCCGCT 562
Qy 382 LysGlyLeuSerTyrLeuHisGluAspCysAsnProLysIleIleHisArgAspIleLys 401
Db 622 AAAGGGCTTGCTTATCTTCATGAGGACTGTCTCTCGTATTATCCATCGCGGATATCAAA 421
Qy 402 AlaSerAsnIleLeuIleAspPheLysPheGluAlaLysValAlaAspPheGlyLeuAla 421
Db 682 GCTGCAATGTCCTTATTGATGATAGCTTCGAGCAAGGTTGCTGATTTGGATTGGCT 682
Qy 422 LysIleAlaSerAspThrAsnThrHisValSerThrArgValMetGlyThrPheGlyTyr 441
Db 742 AAGTTGACTACTGATAATAATACTACTCATGTATCGACTCGTGTCTGGGAACTTTCCGGGTAC 742
Qy 442 LeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPheSerPhe 461
Db 802 CTAGCCCTCGAATATGATCAAGTGGAAAAATTGACAGAGAAGTCTGACGTTTTCTCATTT 802
Qy 462 GlyValValLeuLeuGluLeuIleThrGlyArgArgProValAspAlaAsnAsnValTyr 481
Db 859 GGAGTCATGCTATTGGAACTCATAACTGGGAAGCGACCTGTTGATCACACAAATGCC--- 859
Qy 482 ValAspAspSerLeuValAspTrpAlaArgProLeuLeuAsnArgAla---SerGluGln 500
Db 919 ATGGACGACAGCCTAGTAGACTGGGCTCGACCCTCTGACTCGTGGACTAGAGGAGGAT 919
Qy 501 GlyAspPheGluGlyLeuAlaAspAlaLysMetAsnAsnGlyTyrAspArgGluGluMet 520
Db 979 GGCAACTTTGGAGAGTTGGTTGATGCATTTTGGAGGAAACTACGATCCTCAAGAACTG 979
Qy 521 AlaArgMetValAlaCysAlaAlaAlaCysValArgHisSerAlaArgArgProArg 540
Db 1039 TCAAGATGGCAGCTTGTGCTGCCGTAGCATTCTGTCATTCTGCCAAAAAACGTCCTCAAAA 1039
Qy 541 MetSerGlnIleValArgAlaLeuGluGlyAsnValSerLeuSerAspLeuAsnGluGly 560
Db 1040 ATGAGCCAGATTGTAAGAATATTGGAAGGAGATGTCTCACTGGATGACTTGAAGAGTGGG 1099
Qy 561 MetArgProGlyGlnSerAsnValTyrSerSerTyrGlySerThrAspTyrAspSer 580
Db 1159 ATTAGCCGGGCGCAAAATGTTGCTTACAACTCTTCATCTAGCTCAGATCAGTATGACACA 1159
Qy 581 SerGlnTyrAsnGluAspMetLysLysPheArgLysMetAlaLeuGlyThrGlnGluTyr 600
Db 1160 ATGCAATACAAATGCTGATATGCAGAAGTTGCAGAAAGGCAGTGTCTTCTAATAGCGAGAA 1219
```

```
Qy 601 AsnAlaThrGlyGluTyrSerAsnProThrSerAspTyrGlyLeuTyrProSerGlySer 620
Db 1220 TTGGCACCAGC-----AGTGGCTCC 1240
Qy 621 SerSerGlu 623
Db 1241 AGTGGTGAA 1249
```

Search completed: April 25, 2004, 02:33:32
Job time : 696 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 12, 2004, 20:49:18 ; Search time 5877 Seconds
(without alignments)
4771.643 Million cell updates/sec

Title: US-10-086-464-2

Perfect score: 3453

Sequence: 1 MSSAPSGTGTGSPSPFNST.....REMEMGKIKRTGGYSGPSL 647

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
-Q=/cgn2/1/USPTO_epool_p/US10086464/runat_11052004_121814_7841/app_query.fasta_1.839
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10086464 @CGN_1_1_5265 @runat_11052004_121814_7841 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

1: gb_ba.*

2: gb_htg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_om.*

21: em_or.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pl.*

26: em_ro.*

27: em_sts.*

28: em_un.*

29: em_vi.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pin.*
35: em_htg_rod.*
36: em_htg_mam.*
37: em_htg_vrt.*
38: em_sy.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3453	100.0	1944	6	AX088876	AX088876 Sequence
2	3453	100.0	1944	6	AX825703	AX825703 Sequence
3	3453	100.0	2189	6	AX825705	AX825705 Sequence
4	3453	100.0	2189	6	AY028699	AY028699 Brassica
5	2936.5	85.0	1959	6	AX825738	AX825738 Sequence
6	2936.5	85.0	1959	6	BT008400	BT008400 Arabidops
7	2936.5	85.0	2098	8	AY128792	AY128792 Arabidops
8	2936.5	85.0	2116	8	AY056788	AY056788 Arabidops
9	2936.5	85.0	2188	8	AY059901	AY059901 Arabidops
10	2936.5	85.0	2190	8	AY093065	AY093065 Arabidops
11	2934.5	85.0	1959	8	BT008409	BT008409 Arabidops
12	2934.5	85.0	2257	8	AF370509	AF370509 Arabidops
13	2885	83.6	2324	8	AY089024	AY089024 Arabidops
14	2622	75.9	79706	8	AB020746	AB020746 Arabidops
15	1841.5	52.3	2307	8	AK103247	AK103247 Oryza sat
16	1821.5	52.8	1812	6	AX825735	AX825735 Sequence
17	1812	52.5	80393	8	AP000382	AP000382 Arabidops
18	1748.5	50.6	1902	6	AX088882	AX088882 Sequence
19	1748.5	50.6	1902	6	AX825712	AX825712 Sequence
20	1748.5	50.6	1939	6	AX088881	AX088881 Sequence
21	1748.5	50.6	1939	6	AX825711	AX825711 Sequence
22	1746	50.6	2013	8	BT005955	BT005955 Arabidops
23	1746	50.6	2320	8	AK118488	AK118488 Arabidops
24	1687	48.9	2025	6	AX088885	AX088885 Sequence
25	1687	48.9	2025	6	AX825715	AX825715 Sequence
26	1687	48.9	2104	6	AX088884	AX088884 Sequence
27	1687	48.9	2104	6	AX825714	AX825714 Sequence
28	1668	48.3	1515	6	AX825736	AX825736 Sequence
29	1637.5	47.4	1488	6	AX825737	AX825737 Sequence
30	1574.5	45.6	1185	8	AY224526	AY224526 Oryza sat
31	1567.5	45.4	2364	8	AK064397	AK064397 Oryza sat
32	1567.5	45.4	2371	8	AK102832	AK102832 Oryza sat
33	1507	43.6	82356	8	AC006135	AC006135 Arabidops
34	1507	43.6	92861	8	AC006439	AC006439 Arabidops
35	1480.5	42.9	125021	8	AC007504	AC007504 Arabidops
36	1479.5	42.8	114367	8	ATF10M10	ATF10M10 Arabidops
37	1479.5	42.8	125502	8	ATT4L20	AL023094 Arabidops
38	1479.5	42.8	195921	8	ATCHRIV81	AL161585 Arabidops
39	1459.5	42.3	2158	8	AY113877	AY113877 Arabidops
40	1459.5	42.3	3064	8	AY035076	AY035076 Arabidops
41	1456	42.2	2046	8	AY113039	AY113039 Arabidops
42	1456	42.2	2616	8	AF424623	AF424623 Arabidops
43	1456	42.2	2735	8	AY075681	AY075681 Arabidops
44	1434.5	41.5	120465	2	AC137001	AC137001 Oryza sat
45	1431	41.4	3070	8	AK110196	AK110196 Oryza sat

ALIGNMENTS

RESULT 1

AX088876
 LOCUS AX088876 1944 bp DNA linear PAT 17-MAR-2001
 DEFINITION Sequence 1 from Patent WO0114563.
 ACCESSION AX088876
 VERSION AX088876.1 GI:13397639
 KEYWORDS

SOURCE Brassica napus (rape)

ORGANISM Brassica napus
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

REFERENCE 1

AUTHORS Goring, D. and Silva, N.

TITLE Proline-rich extensin-like receptor kinases

JOURNAL Patent: WO 0114563-A 1 01-MAR-2001;

Goring, Daphne (CA) ; Silva, Nancy (CA)

FEATURES Location/Qualifiers

source

1..1944

/organism="Brassica napus"

/mol_type="unassigned DNA"

/db_xref="taxon:3708"

1..1944

/note="unnamed protein product"

/codon_start=1

/protein_id="CAC34703.1"

/db_xref="GI:13397640"

/translation="MSSAPSGTGTGSPSPSNTSTTTPPASAPPTTPSPSPPTSTI

PTSPSPSRSTPSAPPSTPTSPSPPLPQSPAPPTPGSPSPAPVPTPTNPP
 SVGPSPNREGSPPSPSPSPSSDGLSTGVVGTAGTGGVALLVITLLCLC
 KKRRDEADYVFPFPGKAGTGGQQQROQNTAPPSPDHTVSLPFPKAP
 SPQPFPPTPPMSSSGSDYSDRVLPPSPGLVGFSGSKFTTVEELARATNGS
 EANLGGGGYVHKGLVPSGKAVKLVKVGQGERFQAEVEIISRVRHRLVSL
 VGYCIAGAKLLVVEFPNNLELHLEGRTPTMEWTRLKIALGSAKGLSYLHEDCN
 PKIIRDKASNLILDFKFAKVADFGLAKIADTNTHVTRVMGTGFLAPEVAAGS
 KLTEKSVFSGVLLLELIGRRPVDANNVVDLSVDWAPLILNRASEQDFGLAD
 AOMNNGIDREEMARWCAACACVRRHSARRRMSQIVRALEGNVSLDLNEGRPPGS
 NVYSYSGSDTDYSSQNDYDMKFRKMAIGTQYNAATGEVSNPTSDYLYSGSSSEG
 QTTREMEMGKIKRTGGYSGPSL"

ORIGIN

Alignment Scores:

Pred. No.: 4,02e-76 Length: 1944
 Score: 3453.00 Matches: 647
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-10-086-464-2 (1-647) x AX088876 (1-1944)

Qy 1 MetSerSerAlaProSerProGlyThrGlySerProProSerProProSerAsnSerThr 20
 Db 1 ATGTCCTCGGCGCGTCTCGGGGACTGGTTCCGCTCCATCTCCACCATCAAACTCCACA 60
 Qy 21 ThrThrThrProProAlaSerAlaProProProThrThrProSerSerProProPro 40
 Db 61 ACCACCACTCTCTCCAGCTTCGGCTCTCTCTCCACCAACACCTTCTCTCTCTCCGCG 120
 Qy 41 ProSerThrThrProProSerProProSerSerSerArgSerThrProSerAlaProPro 60
 Db 121 CCATCCCACTATTCGACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180
 Qy 61 ProSerProProThrProSerThrProGlySerProProProProLeuProGlnProSerPro 80
 Db 181 CCATCTCCACCACTCCATCTACGCGGGATCTCCACCTCTCTCTCTCTCTCTCTCTCTCT 240
 Qy 81 ProAlaProThrThrProGlySerProProAlaProValThrProProThrArgAsnPro 100
 Db 241 CCCGCTCCCACTACGCGCGGATCTCCACCCGACCTGTGTACTCTCTCTCTCTCTCTCTCT 300
 Qy 101 ProProSerValProGlyProProSerAsnProSerArgGluGlyGlySerProArgPro 120

Db 301 CCACCTTCAGTCCAGGACCACCGTCCAACTCTTCACGCGAAGGAGATCTCCTCGACCT 360
 Qy 121 ProSerSerProProSerProSerSerSerSerSerSerSerSerSerSerSerSerSerVal 140
 Db 361 CCATCTCTCTCTCTCGCGCGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 420
 Qy 141 GlyTleAlaIleGlyGlyValAlaLeuLeuValIleValThrLeuIleCysLeuLeuCys 160
 Db 421 GGAATCCCACTCGGAGGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 480
 Qy 161 LysAlaLysArgArgArgAspGluGluAspAlaTyrTyrValProProProProProPro 180
 Db 481 AAGAAGAAACGACGAG 540
 Qy 181 GlyProLysAlaGlyGlyProTyrGlyGlyGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200
 Db 541 GGTCCCAAAACGCGGAGGACCTTACGGTGGACAGCAACAATGGCGGCAACAAACGCA 600
 Qy 201 ThrProProSerAspHisValValThrSerLeuProProProProProProProProPro 220
 Db 601 ACACCACTCAGATCATGTCGTGACGTCACTACCAACCAACCACTAAGGCTTCACTCCA 660
 Qy 221 ProArgGlnProProProProProProProProProProProProProProProProPro 240
 Db 661 CCACGGCAACTCTCTCCACTCCACCGCTTTCATGACACAGCGGCGGCTCCGAC 720
 Qy 241 TyrSerAspArgProValLeuProProProProProProProProProProProProPro 260
 Db 721 TACTCGGACCGTCAGTCT 780
 Qy 261 SerThrPheThrTyrGluGluLeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeu 280
 Db 781 AGCATTTTCATACAGAGAGAGTACGTAGAGCCACCAATGGTCTTCTCGAGCGCAACTG 840
 Qy 281 LeuGlyGlnGlyGlyPheGlyTyrValHisGlyValLeuProSerGlyLysGluVal 300
 Db 841 TTAGGACAGCGCGGTTCCGTTACCTGCACAAAGGTGTGTTCCTAGTGGGAAGAAGTT 900
 Qy 301 AlaValLysGlnLeuLysValGlySerGlyGlnGlyGluArgGluPheGlnAlaGluVal 320
 Db 901 GCTGTGAAGCAGTTGAAGTTGGAGTGGTCAGGAGAGAGAGAGTTCAGGACAGAGTT 960
 Qy 321 GluIleIleSerArgValHisHisArgHisLeuValSerLeuValGlyTyrCysIleAla 340
 Db 961 GAGATCATACAGAGAGTTCAACACAGAGATCTGGTGTCTCTTGTGGTTATGTGATCGCC 1020
 Qy 341 GlyAlaLysArgLeuLeuValTyrGluPheValProAsnAsnAsnLeuGluLeuHisLeu 360
 Db 1021 GGTGCCAAAGATTGCTTGTCTATGAGTTTGTCTTAAACAACATCTCGAGCTTCACCTC 1080
 Qy 361 HisGlyGluGlyArgProThrMetGluTrpSerThrArgLeuLysIleAlaLeuGlySer 380
 Db 1081 CATGCGGAGGACGGCTTCAATGGAATGGAGCACCAAGATTGAAGATTGCTCTTTGGATCT 1140
 Qy 381 AlalysGlyLeuSerTyrLeuHisGluAspCysAsnProLysIleIleHisArgAspIle 400
 Db 1141 GCTAAAGAGCTTCTTATCTTCATGAAGATTGCAATCTCTAAATCATTTACCGGTATATC 1200
 Qy 401 LysAlaSerAsnIleLeuIleAspPheLysPheGluAlaLysValAlaAspPheGlyLeu 420
 Db 1201 AAGCTTCAAAACATATTGATAGATTTCAGTTTGAAGCTTAAGTTGCTGATTTTGGTCT 1260
 Qy 421 AlalysIleAlaSerAspThrAsnThrHisValSerThrArgValMetGlyThrPheGly 440
 Db 1261 GCTAAGATTGCTCTGATACAAACACGATGATCAACACGATGATCAACGATGATGGAACCTTTGG 1320
 Qy 441 TyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPheSer 460
 Db 1321 TACTTGGCTCGGAATACGCTGCAAGCGGAAGCTCACGGAGAAGTCTGACGTTTCTTCA 1380
 Qy 461 PheGlyValValLeuLeuGluLeuIleThrGlyArgArgProValAspAlaAsnVal 480
 Db 1381 TTTGGCGTTGCTTTTGGAGCTCATTTACTGGACGTCGACCGCTTGATGCCCAACAATGTC 1440

Qy 481 TyrValAspSerLeuValAspTyrAlaArgProLeuLeuAsnArgAlaSerGluGln 500
 Db 1441 TATGTAGATGACAGCTAGTTGACTGGGCGACGACCATTCCTTAACCGGACATCTGACGAA 1500

Qy 501 GlyAspPheGluGlyLeuAlaAspAlaLysMetAsnGlyTyrAspArgGluGluMet 520
 Db 1501 GGAGACTTTGAGGGTTAGCTGATGCCAAGATGAATAATGGGTATGACAGAGAGGATG 1560

Qy 521 AlaArgMetValAlaCysAlaAlaCysValArgHisSerAlaArgAtqProArg 540
 Db 1561 GCTCGCATGGTGTGTGCTGGCGCTGTGTTCCGCAATCAGCTCGCGCGACCTCGC 1620

Qy 541 MetSerGlnIleValArgAlaLeuGluGlyAsnValSerLeuSerAspLeuAsnGluGly 560
 Db 1621 ATGAGCCAGATTGTCGGTGTAGNAGGAATGATACATCTGTCAGATCTTAACGAAGG 1680

Qy 561 MetArgProGlyGlnSerAsnValTyrSerSerTyrGlyGlySerThrAspTyrAspSer 580
 Db 1681 ATGAGACAGGTCAAAGCAATGTATACAGCTCATACGAGGAGGACCGATTAAGACTCG 1740

Qy 581 SerGlnTyrAsnGluAspMetLysLysPheArgLysMetAlaLeuGlyThrGlnGluTyr 600
 Db 1741 AGCCAGTACATGAAGACATGAAGAAGTTTAGGAAAATGGCACCTGGAACCTCAAGAGTAC 1800

Qy 601 AsnAlaThrGlyGluTyrSerAsnProThrSerAspTyrGlyLeuTyrProSerGlySer 620
 Db 1801 AACGCCAGGTGAGTACAGTAATCCGACCACTGACTATGACTGTACCCGCTCGGTCA 1860

Qy 621 SerSerGluGlyGlnThrThrArgGluMetGluMetGlyLysLysLysArgThrGlyGln 640
 Db 1861 AGCAGCGAGGGCCAAACCAACGCAATGGAGATGGGAAGATTAAGAGAACCGGTGAG 1920

Qy 641 GlyTyrSerGlyProSerLeu 647
 Db 1921 GGTTATAGTGACCTTCTCTT 1941

RESULT 2
 AX825703
 LOCUS AX825703 1944 bp DNA linear PAT 11-DEC-2003
 DEFINITION Sequence 1 from Patent WO03072763.
 ACCESSION AX825703
 VERSION AX825703.1 GI:39751230
 KEYWORDS
 SOURCE Brassica napus (rape)
 ORGANISM Brassica napus
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
 1
 Goring, D., Silva, N. and Haffani, Y. Z.
 Increasing plant seed production
 Patent: WO 03072763-A 1 04-SEP-2003;
 Goring, Daphne (CA) ; Silva, Nancy (CA) ; Haffani, Yosr, Z. (CA)
 Location/Qualifiers
 1..1944
 /organism="Brassica napus"
 /mol_type="unassigned DNA"
 /db_xref="taxon:3708"
 1..1944
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAE84846.1"
 /db_xref="GI:39751231"
 /translation="MSSAPSGTGSPPSPSNSTTTTPPPASAPPPPTPPSPPPSTT
 PTPSPSPSRSTPSAPPSPPTPSPSPPLPQSPSPAPPTPGSPSPAPPTPPTRNPP
 VPGPPSPNPSRGSGSPSPSPSPSDGLSTGVVVGIAIGGVALLVITVLICLIC
 KKRRRDEDAYVPPPPPGPKAGGPGYGOQOORONATPPSDHVVTLGTPPKAP
 SPPOPPPPPPPMSPSSSGSDYSDRVLPPSPGLVGLSKSTFTVEELARATNGS
 EANLGGGFGYVHKVLPSGKEVAVKQLKVGSCQGERFEQAEVILISRVHRLHLSL
 VGYCAGAKLLVTFVFNNNLEHLHGEGRPTMSEWTRLKIALGASGLSYLHEDCN
 PKIHRDKASNLILIDFKFAKVADFLAKIASDTNTHVSTVMTGTFGLAPEYASG
 KLTEKSDVFGVGLLELITGRRPVDANNVVDLSLVDWARPLNLRASEQGDPEGLAD

AKNNNGYDREEMARMVACAAACVRHSARRRPMRSQIVRALEGNVLSLSDLEGMRRPGOS
 NVTSSYGGSTDYDSOQVEDMKPRKMLGTQBYNATGEYSNPTSDYGLYSGSSESG
 QTTREMEMGKI KRTGGYSGPSL"

ORIGIN

Alignment Scores:
 Pred. No.: 4.02e-76 Length: 1944
 Score: 3453.00 Matches: 647
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-10-086-464-2 (1-647) x AX825703 (1-1944)

Qy 1 MetSerSerAlaProSerProGlyThrGlySerProProSerProSerAsnSerThr 20
 Db 1 ATGTCTCTCGGGCGCGTCTCCGGGAGCTGGTTCGGCTCCATCTCCACCATCAAACTCCACA 60

Qy 21 ThrThrThrProProAlaSerAlaProProProThrThrProSerSerProProPro 40
 Db 61 ACCACCACTCTCTCCAGCTTCGGCTCCTCCACACACACCTTCTCTCTCCGCG 120

Qy 41 ProSerThrIleProThrSerProProProSerSerArgSerThrProSerAlaProPro 60
 Db 121 CCATCCACTATTCGACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180

Qy 61 ProSerProProThrProSerThrProGlySerProProProLeuProGlnProSerPro 80
 Db 181 CCATCTCCAACTCCATCTACGCGGAGATCTCCACCTCTCTCTCTCTCTCTCTCTCTCA 240

Qy 81 ProAlaProThrThrProGlySerProProAlaProValThrProProThrArgAsnPro 100
 Db 241 CCGCTCCAATACTAGCGCGGATCTCCACCGCGACCTGTACTCTCTCTCTCTCTCTCTCT 300

Qy 101 ProProSerValProGlyProProProSerAsnProSerArgGluGlyGlySerProArgPro 120
 Db 301 CCACCTTCAGTCCGAGGACCGGTCCAAATCTCTTCGCGAAGGAGGATCTCTCTCGACCT 360

Qy 121 ProSerSerProProProProProSerProSerSerAspGlyLeuSerThrGlyValValVal 140
 Db 361 CCATCT 420

Qy 141 GlyIleAlaIleGlyGlyValAlaLeuLeuValIleValThrLeuIleCysLeuLeuCys 160
 Db 421 GGAATCGCCATCGGAGGAGTCT 480

Qy 161 LysLysLysArgArgArgAspGluAspAlaTyrTyrValProProProProProProPro 180
 Db 481 AAGAAGAAACGCGGAGAGACGAAAGATGCTTACTATGTTCTCTCTCTCTCTCTCTCTCTCT 540

Qy 181 GlyProLysAlaGlyGlyProTyrGlyGlyGlnGlnGlnTyrArgGlnGlnAsnAla 200
 Db 541 GGTCCCAAAACCGGAGGACCTTACGGTGGAGCAGCAGCAACAATGGCGGCAACAAACGCA 600

Qy 201 ThrProProSerAspHisValValThrSerLeuProProProProLysAlaProSerPro 220
 Db 601 ACACCACTCGATCATGTCGTGAGCTACTACCAACCAACCACTTAAGGCTCTCATCTCCA 660

Qy 221 ProArgGlnProProProProProProProPheMetSerSerSerGlyGlySerAsp 240
 Db 661 CCACGGCAACTCTCTCCACCTCCACCGCTTTCATGAGCAGCAGCGCGCTCCGAC 720

Qy 241 TyrSerAspArgProValLeuProProProSerProGlyLeuValLeuGlyPheSerLys 260
 Db 721 TACTCGACCGCTCAGTTCCTCTCCACCGCTCTCCAGGGCTTGTGTAGGCTTCTCCAAA 780

Qy 261 SerThrPheThrTyrGluGluLeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeu 280
 Db 781 AGCACTTTCATACGAGGAGCTAGCTAGACCAACCACTGTTCTCTCCGAGGCAACTTG 840

Qy 281 LeuGlyGlnGlyPheGlyTyrValHisLysGlyValLeuProSerGlyLysGluVal 300

Db 841 TTAGACAAGCGGGTTTCGGTTACGTGCACAAAGGTGTGTTGCCCTAGTGGGAAGAAGTT 900
Qy 301 AlaVallysGlnLeuLysValGlySerGlyGlnGlyGluArgGluPheGlnAlaGluVal 320
Db 901 GCTGTGAAGCAGTTGAAAGTTGGGAGTGGTCAGGGAGAGAGAGTTTCAGCCAGAGGTT 960
Qy 321 GluIleIleSerArgValHisArgHisIleuValSerLeuValGlyTyrCysIleAla 340
Db 961 GAGATCATCAGCAGAGTTTACCACACAGGATCTGGTGTCTCTTGTGGTTATTGATCGCC 1020
Qy 341 GlyAlaLysArgLeuLeuValTyrGluPheValProAsnAsnLeuGluLeuHisLeu 360
Db 1021 GGTGCCAAAAGATTGCTTGTCTATGAGTTTGTCTTAACAACAATCTCGAGCTTCACCTC 1080
Qy 361 HisGlyGluGlyArgProThrMetGluTrpSerThrArgLeuLysIleAlaLeuGlySer 380
Db 1081 CATGCGAGGAGCGGCTTACAATGGAATGGAGCACCAGATTGAAGATTGCTCTTGGATCT 1140
Qy 381 AlalysGlyLeuSerTyrLeuHisGluAspCysAsnProLysIleIleHisArgAspIle 400
Db 1141 GCTAAAGGACTTCTTATCTTCATGAAGATTGCAATCTCAAAATCATTCACCGTGATATC 1200
Qy 401 LysAlaSerAsnIleLeuIleAspPheLysPheGluAlaLysValAlaAspPheGlyLeu 420
Db 1201 AAGGCTTCAACATATTGATAGATTCAAGTTTGAAGCTAAGTTGCTGATTTGGTCTT 1260
Qy 421 AlalysIleAlaSerAspThrAsnThrHisValSerThrArgValMetGlyThrPheGly 440
Db 1261 GCTAAGATTGCTTCTGATACAAACACGCGATGTATCAACACGCTGTATGGGAACCTTTGGG 1320
Qy 441 TyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPheSer 460
Db 1321 TACTTGGCTCCGGAATACGCTGCAAGCGGAAGCTCACGGAGAAGTCTGACGCTTTCTCA 1380
Qy 461 PheGlyValValLeuLeuGluLeuIleThrGlyArgArgProValAspAlaAsnVal 480
Db 1381 TTTGGCGTTGTCTTTTGAGCTCATTTACTGAGCTGCACCGCTGTATGCCAAACAATGTC 1440
Qy 481 TyrValAspAspSerLeuValAspTrpAlaArgProLeuLeuAsnArgAlaSerGluGln 500
Db 1441 TATGTAGATGACAGCTTATGTTGACTGGGCACGACCATTTGCTTAACCGAGCATCTGAGCA 1500
Qy 501 GlyAspPheGluGlyLeuAlaAspAlalysMetAsnAsnGlyTyrAspArgGluMet 520
Db 1501 GGAGACTTTGAGGGTTTGTAGCTGATGCAAAAGATGAATAATGGGTATGACAGAGAGAGATG 1560
Qy 521 AlaArgMetValAlaCysAlaAlaCysValArgHisSerAlaArgArgProArg 540
Db 1561 GCTGCAATGGTTGCTTGTGCTGGGCTTGTGTTGCCCATTCAGCTCGCGCGAGACCTGCG 1620
Qy 541 MetSerGlnIleValArgAlaLeuGluGlyAsnValSerLeuSerAspLeuAsnGluGly 560
Db 1621 ATGACCCAGATTGTGCGTGGCTTAGAAGAAATGTATCACTGTCAAGATCTTAAACGAAGG 1680
Qy 561 MetArgProGlyGlnSerAsnValTyrSerSerTyrGlyGlySerThrAspTyrAspSer 580
Db 1681 ATGACCCAGGTCAAAGCAATGTATACAGCTCATACGGAGGAAGCACCAGATTATGACTCG 1740
Qy 581 SerGlnTyrAsnGluAspMetLysLysPheArgLysMetAlaLeuGlyThrGlnGluTyr 600
Db 1741 AGCCNGTACATGAGACATGAGAGATTAGGAATTTAGGAATTTGGCACTTGGAACTCAAGATAC 1800
Qy 601 AsnAlaThrGlyGluTyrSerAsnProThrSerAspTyrGlyLeuTyrProSerGlySer 620
Db 1801 AACGCCAGGGTGAGTACAGTAATCCGACCAGTGACTATGGACTGTACCCGCTGTTTCA 1860
Qy 621 SerSerGluGlyGlnThrArgGluMetGluMetGlyLysIleLysArgThrGlyGln 640
Db 1861 AGCAGCGGGGCCAAACCCACACCGGAAATGGAGATGGGGAAGATTAAAGAAACCGGTGAG 1920
Qy 641 GlyTyrSerGlyProSerLeu 647
Db 1921 GGTATATAGTGACCTTCTCTT 1941

RESULT 3

AX825705

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

ORIGIN

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

US-10-086-464-2 (1-647) x AX825705 (1-2189)

Qy 1 MetSerSerAlaProSerProGlyThrGlySerProProSerProProSerAsnSerThr 20

Db 97 ATGTCTTCGGCGCGCTCTTCGGGGAGCTGGTTCGGCTTCACTCCACCATCAAACTCCACA 156

Qy 21 ThrThrThrProProAlaSerAlaProProProThrThrProSerSerProProPro 40

Db 157 ACCACCACTCTCTCCAGCTTCCGCTCTCTCTCCACCAACACCTTCTCTCTCCGCGC 216

Qy 41 ProSerThrIleProThrSerProProSerSerSerArgSerThrProSerAlaProPro 60

Db 217 CCATCCACTATTCGACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 276

Qy 61 ProSerProProThrProSerThrProGlySerProProProProProGlnProSerPro 80

Db 277 CCATCTCCACCACTCATCTACCGCGGATCTCCACCTCTCTCTCTCTCTCTCTCTCTCA 336

Qy 81 ProAlaProThrThrProGlySerProProAlaProValThrProProThrArgAsnPro 100

Db 337 CCGCTCCAACTACGCGCGGATCTCCACCGGACCTGTTACTCTCTCTCTCTCTCTCTCTCT 396

Qy 101 ProProSerValProGlyProProProSerAsnProSerArgGluGlyGlySerProArgPro 120

Db 397 CCACCTTCAGTCCAGGACCACTCCCAATCTTTCGCGAAGGAGGATCTCTCTCTCTCTCT 456

Qy 121 ProSerSerProProProProSerProSerSerAspGlyLeuSerThrGlyValValVal 140

Db 457 CCATCT 516

Qy 141 GlyIleAlaIleGlyValAlaLeuLeuValIleValThrLeuIleCysLeuLeuCys 160

Db 517 GGAATCGCATCGGAGGAGTCT 576

Qy 161 LysIleLysArgArgArgAspGluAspAlaTyrTyrValProProProProProProPro 180

Db 577 AAGAAGAAACACGCGAGAGACGAAGAAGATGCTTACTATGTCTCTCTCTCTCTCTCTCTCT 636

Qy 181 GlyProLysAlaGlyGlyProTyrGlyGlyGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200

AX825705
Sequence 3 from Patent WO03072763.
AX825705
AX825705.1 GI:39751232

2189 bp

DNA

linear

PAT 11-DEC-2003

Brassica napus (rape)
Brassica napus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.1
Goring, D., Silva, N. and Haffani, Y. Z.

Increasing plant seed production

Patent: WO 03072763-A 3 04-SEP-2003;

Goring, Daphne (CA) ; Silva, Nancy (CA) ; Haffani, Yosr, Z. (CA)

Location/Qualifiers

1. .2189

/organism="Brassica napus"

/mol_type="unassigned DNA"

/db_xref="taxon:3708"

4.38e-76

Length: 2189

Matches: 647

Conservative: 0

Mismatch: 0

Indels: 0

Gaps: 0

DB:	8	Gaps:	0
US-10-086-464-2 (1-647) x AY028699 (1-2189)			
QY	1	MetSerSerAlaProSerProGlyThrGlySerProProSerProProSerAsnSerThr	20
DB	97	ATGTCCTCGGCGCGCTCTCCGGGACTGGTTGCGCTCCATCTCCACCATCAAACTCCACA	156
QY	21	ThrThrThrProProAlaSerAlaProProProThrThrProProSerSerProProPro	40
DB	157	ACCACCACTCCTCCAGCTTCGGCTCCTCTCCACACACACCTTCTCTCTCCGCGC	216
QY	41	ProSerThrIleProThrSerProProSerSerArgSerThrProSerAlaProPro	60
DB	217	CCATCCACATATTCGACATCTCCCTCCTCTCTCTCCGCTCTACACCTTCCTCCTCCT	276
QY	61	ProSerProProThrProSerThrProGlySerProProProLeuProGlnProSerPro	80
DB	277	CCATCTCCACCACTCCATCTACGCCGGATCTCCACCTCTCTCTCCAGCCGCTCTCCA	336
QY	81	ProAlaProThrThrProGlySerProProAlaProValThrProProThrArgAsnPro	100
DB	337	CCCGCTCCAACTAGCCCGGATCTCCACCACCTGTACTCTCTCTACTCTCGAAACCT	396
QY	101	ProProSerValProGlyProProProSerAsnProSerArgGluGlyGlySerProArgPro	120
DB	397	CCACCTTCAGTCCCGAGGACCAACCGTCCCAATCTTACCGGAAGGAGGATCTCTCGACCT	456
QY	121	ProSerSerProSerProProSerProSerSerAspGlyLeuSerThrGlyValVal	140
DB	457	CCATCTCTCCTCGCCCGCTCTCCTCTCTCCGACGGTTTATCAACAGGAGTGGTGGTG	516
QY	141	GlyIleAlaIleGlyGlyValAlaLeuLeuValIleValThrLeuIleCysLeuLeuCys	160
DB	517	GGAAATCGCCATCGGAGGAGTGGCTGCTGTGTGATAGTACTCTGATTTGCTCTCTGT	576
QY	161	LysIleValArgArgAspGluAspAlaTyrTyrValProProProProProProPro	180
DB	577	AAGAAGAAACACGCGAGAGACGAAAGATGCTTACTATGTTCCTCCGCCACCTCTCTCT	636
QY	181	GlyProLysAlaGlyGlyProTyrGlyGlyGlnGlnGlnTyrArgGlnGlnAlaAsnAla	200
DB	637	GGTCCCAAGCGGNGGACCTTACGGTGGACAGCAGCAACAATGGCGGCACACAAACGCA	696
QY	201	ThrProProSerAspHisValThrSerLeuProProProProProProProProPro	220
DB	697	ACACCACCGTCAGATCATGTGTCGTCAGCTCACTTACCACCAACCACTTAAGGCTCCATCTCCA	756
QY	221	ProArgGlnProProProProProProProPheMetSerSerSerGlyGlySerAsp	240
DB	757	CCACGGCAACCTCTCTCCACCTCCACCGCTTTTCATGAGCAGCAGCGCGGCTCCGAC	816
QY	241	TyrSerAspArgProValLeuProProProSerProGlyLeuValLeuGlyPheSerLys	260
DB	817	TACTCGGACCGTCCAGTCTTCTCTCCACCGCTCTCCAGGGCTTGTGTAGGCTTCTCCAAA	876
QY	261	SerThrPheThrTyrGluGluLeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeu	280
DB	877	AGCACTTTACATACAGGAGCTAGCTAGAGCCACCAATGGTTTCTCCGAGCGCGAACTTG	936
QY	281	LeuGlyGlnGlyGlyPheGlyTyrValHisLysGlyValLeuProSerGlyLysGluVal	300
DB	937	TTAGGACAAGCGGGTTCGGTTACGTGCACAAAGGTGTGTTCCTTAGTGGGAAAGAGGT	996
QY	301	AlaValLysGlnLeuLysValGlySerGlyGlnGlyGluArgGluPheGlnAlaGluVal	320
DB	997	GCTGTGAACAGCTGAAAGTTGGGAGTGGTTCAGGGAGAGAGGAGTTTCAGCGAGAGGT	1056
QY	321	GluIleIleSerArgValHisArgHisLeuValSerLeuValGlyTyrCysIleAla	340
DB	1057	GAGATCATCAGCAGAGTTCACACAGGCACTCTGGTGTCTCTTGTGGTTATTGATCGCC	1116
QY	341	GlyAlaLysArgLeuLeuValTyrGluPheValProAsnAsnLeuLeuLeuHisLeu	360

DB	1117	GGTCCAAAAGATTGCTTGCTCTATGAGTTGTTCTTAACAACAATCTCGAGCTTCACCTC	1176
QY	361	HisGlyGluGlyArgProThrMetGluTrpSerThrArgLeuLysIleAlaLeuGlySer	380
DB	1177	CATGCGGAGGACGGCTCACAATGGAATGGAGCACCAAGATTGAAGATTGCTCTTGGATCT	1236
QY	381	AlalysGlyLeuSerTyrLeuHisGluAspCysAsnProLysIleIleHisArgAspIle	400
DB	1237	GCTAAAGGACTTCTTATCTTCATGAAGATTGCAATCTCTAAATCATTCACCGTGATATC	1296
QY	401	LysAlaSerAsnIleLeuIleAspPheLysPheGluAlaLysValAlaAspPheGlyLeu	420
DB	1297	AAGCTTCAACAATATTGATAGATTCAAGTTGAAGTGAAGTTGCTGATTTGGTCTT	1356
QY	421	AlalysIleAlaSerAspThrAsnThrHisValSerThrArgValMetGlyThrPheGly	440
DB	1357	GCTAAGATTGCTTCTGATACAAACACGATGATCAACACGCTGTGATGGGAACCTTTGGG	1416
QY	441	TyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPheSer	460
DB	1417	TACTTGGCTCGGAATACGCTGCAAGCGGAAAGCTCACGGAGAAGTCTGACGTTTCTCA	1476
QY	461	PheGlyValValLeuLeuLeuIleThrGlyArgArgProValAspAlaAsnVal	480
DB	1477	TTTGGCGTTGCTTTTGGAGCTCATCTAGGACGTGACCCGTTGATGCCACAATGTC	1536
QY	481	TyrValAspSerLeuValAspTrpAlaArgProLeuLeuAsnArgAlaSerGluGln	500
DB	1537	TATGTAGATGACAGCTTAGTTGACTGGGCACGACCATTTGCTTAACCGAGCATCTGAGCAA	1596
QY	501	GlyAspPheGluGlyLeuAlaAspAlaLysMetAsnAsnGlyTyrAspArgGluGluMet	520
DB	1597	GGAGACTTTGAGGGTTAGCTGATGCAAGAGATGAATAATGGGTATGACAGAGGAGATG	1656
QY	521	AlaArgMetValAlaCysAlaAlaAalaCysValArgHisSerAlaArgArgProArg	540
DB	1657	GCTCGCATGGTTGCTTGTGCTGGCGCTTGTGTTCCCATTCAGCTCGCCGAGACCTCGC	1716
QY	541	MetSerGlnIleValArgAlaLeuGluGlyAsnValSerLeuSerAspLeuAsnGluGly	560
DB	1717	ATGAGCCAGATTGTCGTGCGTTAGAGGAAATGTATCACTGTCTAGATCTTAACGAAGGG	1776
QY	561	MetArgProGlyGlnSerAsnValTyrSerSerTyrGlyGlySerThrAspTyrAspSer	580
DB	1777	ATGAGCAGGCTCAAGCAATGTATACAGCTCATACGGAGGAGACCGATTATGACTCG	1836
QY	581	SerGlnTyrAsnGluAspMetLysPheArgLysMetAlaLeuGlyThrGlnGluTyr	600
DB	1837	AGCCAGTACATGAAGACATGAAGAGTTTAGGAAAATGGCACTTGGNACTCAAGAGTAC	1896
QY	601	AsnAlaThrGlyGlyTyrSerAsnProThrSerAspTyrGlyLeuTyrProSerGlySer	620
DB	1897	AACGCCACGGGTAGTACAGTAATCCGACCACTATGAGTGTATGACCTGATCCCTCTGGTTCA	1956
QY	621	SerSerGluGlyGlnThrThrArgGluMetGluMetGlyLysIleLysArgThrGlyGln	640
DB	1957	AGCAGCAGGCGCCAAACACACACCGGAAATGGAGATGGGAAAGATTAAAGAAACCGGTGAG	2016
QY	641	GlyTyrSerGlyProSerLeu	647
DB	2017	GGTTATAGTGACCTTCTCTT	2037

RESULT 5

AX825738	AX825738	1959 bp	DNA	linear	PAT 11-DEC-2003
LOCUS					
DEFINITION	Sequence 36 from Patent WO03072763.				
ACCESSION	AX825738				
VERSION	AX825738.1	GI:39751255			
KEYWORDS	Arabidopsis thaliana (thale cress)				
SOURCE	Arabidopsis thaliana				
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE

1 Goring, D., Silva, N. and Haffani, Y. Z.

Increasing plant seed production

Patent: WO 03072763-A 36 04-SEP-2003;

JOURNAL Daphne (CA); Silva, Nancy (CA); Haffani, Yosr, Z. (CA)

Goring, Location/Qualifiers

FEATURES

1. 1959

source /organism="Arabidopsis thaliana"

/mol_type="unassigned DNA"

/db_xref="taxon:3702"

ORIGIN

Alignment Scores:

Pred. No.: 1,23e-63 Length: 1959
Score: 2936.50 Matches: 574
Percent Similarity: 90.61% Conservative: 24
Best Local Similarity: 86.97% Mismatches: 41
Query Match: 85.04% Indels: 21
DB: 6 Gaps: 12

US-10-086-464-2 (1-647) x AX825738 (1-1959)

Qy 1 MetSerSerAlaProSerProGlyThr-----GlySerProProSerProProSer 17
Db 1 ATGTCACAGCGCGTCTCCAGGCACTACTCCATCACCATCTCCACCGTCTCTCCACACA 60
Qy 18 AsnSerThrThrThrProProProAla-----SerAlaProProThrThrPro 35
Db 61 AACTCGCAACACCACTCTCTCCAGCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 120
Qy 36 SerSerProProPro-----ProSerThrThrThrThrThrThrThrThrThrThr 54
Db 121 TCTTCT 180
Qy 55 ThrProSerAlaProProProSerProProThrThrProThrThrProGlySerProPro 74
Db 181 CCT 228
Qy 75 LeuProGlnProSerProProAlaProThrThrThrProGlySerProProAlaProVal 94
Db 229 CTTCCTCAACCTTCCCTCCGCTCCCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 285
Qy 95 ProProThr-----ArgAsnProProProSerValProGlyProProSerAsnProSer 113
Db 286 CCCTCAACCTCTGAAGCCCTCTCATCTCTCTTAACCAAGGACCAACACCTCTCTCTCA 342
Qy 114 GluGlyGlySerProArgProProSerSer-----ProSerProProSerProSerSer 131
Db 343 ---GGATCTACTCTAGAACCTCCATCAACACTAAACCGTCCGCGCTCTGATCTCTCC 399
Qy 132 AspGlyLeuSerThrGlyValValValGlyIleAlaIleGlyGlyValAlaLeuLeuVal 151
Db 400 GATCGATTGTCTACCGGAGTTGTGTAGGATCGCCATTTGGAGGAGTCTATTTCTTGT 459
Qy 152 IleValThrIleLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 169
Db 460 ATACTGACTCTGATTGTCTCTCTCTGTAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 519
Qy 170 AspAlaTyrTyrValProProProProProGlyProGlyProGlyProGlyProGly 189
Db 520 GCTGCTTACTATGTTCT 579
Qy 190 GlyGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 208
Db 580 GGTCAACAACTAGTATTGGCAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 639
Qy 209 ThrSerLeuProProProProProProProProProProProProProProProProPro 228
Db 640 ACGTCATTG-----CCACCACTTAAGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 696
Qy 229 ProProPro-----PheMetSerSerSerGlyGlySerAspTyrSerAspArgProVal 247

Db 697 CCACCACGACGATTCATGAGTAGCAGTGGTGTCTGACTATTTCGATCTTCCGGTCTCT 756
Qy 248 ProProProSerProGlyLeuValLeuGlyPheSerLySerThrPheThrTyGluGlu 267
Db 757 CTTCCACCATCTCCAGGCTTGTGTAGGCTTTTCTAAAAGCACTTTTACATATGAGGAG 816
Qy 268 LeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyPheGly 287
Db 817 TTGTCGAGAGCTACTATGCTCTCTGAGGCTAATTTGTAGGCAAGGAGGTTTGGT 876
Qy 288 TyrValHisLyGlyValLeuProSerGlyLyGlyValAlaValLyGlnLeuLyVal 307
Db 877 TATGTGCATAAAGGTATATTGCTAGTGGGAAAGAGTGTCTGTGAACAGTTCGAAAGCT 936
Qy 308 GlySerGlyGlnGlyGluArgGluPheGlnAlaGluValGluIleLeuSerArgValHis 327
Db 937 GGTAGTGGTCAGGAGAGAGAGAGTTCAGGCTGAGGTGTGAGATCATTTAGCAGAGTTCAT 996
Qy 328 HisArgHisLeuValSerLeuValGlyTyrCysIleAlaGlyAlaLyAsArgLeuVal 347
Db 997 CACAGGCAATTTGGTTCTCTTATTGTTATGATGCGCGGTGTTCAAGATTTACTTGT 1056
Qy 348 TyrGluPheValProAsnAsnAsnLeuGluLeuHisGlyGlyGlyGlyArgProThr 367
Db 1057 TATGAGTTTGTTCACAAACAATCTTGTAGTTTCCCTCCATGTTAAGGAGCGCTACG 1116
Qy 368 MetGluTrpSerThrArgLeuLyIleAlaLeuGlySerAlaLyGlyLeuSerTyLeu 387
Db 1117 ATGGAATGGAGTACTAGATTGAAGATGCTCTTGGATCTGCTAAAGGACTTTTATATCTT 1176
Qy 388 HisGluAspCysAsnProLyIleIleHisArgAspIleLyAsAlaSerAsnIleLeu 407
Db 1177 CATGAAGATTGCAATCCGAAATCATTCACCGTGATTAAGGCGTCAACATATTGATT 1236
Qy 408 AspPheLyGlyPheGluAlaLyValAlaAspPheGlyLeuAlaLyIleAlaSerAspThr 427
Db 1237 GATTTCAAATTTGAAGTAAAGTGTCTGACTTTTGGTCTTGCAAGATTTGCTTCTGATACA 1296
Qy 428 AsnThrHisValSerThrArgValMetGlyThrPheGlyTyrLeuAlaProGlyTyrAla 447
Db 1297 AACACTCATGTATCTACACGCGTGATGGGAACCTTTTGGGTATTTGGCTCCGGAATAGCT 1356
Qy 448 AlaSerGlyLyLeuThrGluLySerAspValPheSerPheGlyValValLeuLeuGlu 467
Db 1357 GCAAGTGAAGAGCTCACAGAAAGTCTGACGTTTCTCATTTGCGGTGTACTTTTGAA 1416
Qy 468 LeuIleThrGlyArgArgProValAspAlaAsnAsnValTyrValAspAspSerLeuVal 487
Db 1417 CTTATTACTGGAGGCGCCCTGTTGATGCGAACAACTGCTATGTATGATGACAGCTTAGTT 1476
Qy 488 AspTrpAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGlyLeuAla 507
Db 1477 GACTGGGCACGACCATTTGCTTAACCGGACATCTGAGGAGGAGATTTTGGAGGTTTGGCT 1536
Qy 508 AspAlaLyGlyMetAsnAsnGlyTyrAspArgGluGluMetAlaArgMetValAlaCysAla 527
Db 1537 GATTCAAGATGGGTAATGAGTATGACAGAGAGAGATGCTCGCATGTTGTTGGCT 1596
Qy 528 AlaAlaCysValArgHisSerAlaArgArgArgProArgMetSerGlnIleValArgAla 547
Db 1597 GCGCTTGTGTTCCCATCTCAGCTCGCGCAGACCTCGCATGAGCCAGATAGTAGCGGCG 1656
Qy 548 LeuGluGlyAsnValSerLeuSerAspLeuAsnGluGlyMetArgProGlyGlnSerAsn 567
Db 1657 TTGAAGGAAATGATGCTGCTGCTGATCTTAAAGAGGAGATGAGACCGGTCACAGCAAC 1716
Qy 568 ValTyrSerSerTyrGlyGlySerThrAspTyrAspSerSerGlnTyrAsnGluAspMet 587
Db 1717 GTATACAGCTCATATGGAGGAGCAGACTATGACACGAGCCCATACACGAGCATATG 1776
Qy 588 LyLyGlyPheArgLyGlyMetAlaLeuGlyThrGlnGlyTyrAsnAlaThrGlyGlyTyrSer 607

Db 1777 ATAAAGTTTAGGAAATCGCTCTTGAACTCAAGNATACGCACACACCGCGAGTACAGT 1836

Qy 608 AnProThrSerAppTyGlyLeuTyProSerGlySerSerGlyGlnThrThr 627

Db 1837 AATCCACCACTGACTACGAGCTGTACCGCTCTGGTTCAAGCAGTGAAGGTCAAGCCACA 1896

Qy 628 ArgGluMetGluMetGlyLeuIleValysArgThrGlyGlnGlyTySerGlyProSerIleu 647

Db 1897 CGAGAAATGGAGATGGGAAGATTAAAGAAACCGGTCAAGGTATTAGTGGACCTCTCTT 1956

RESULT 6

BT008400

LOCUS

DEFINITION

Arabidopsis thaliana At3g24550 gene, complete cds.

BT008400

VERSION

BT008400.1

KEYWORDS

FLI CDNA.

SOURCE

ORGANISM

Arabidopsis thaliana (thale cress)

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids II; Brassicales; Brassicaceae; Arabidopses.

1 (bases 1 to 1959)

Chen, R., Chen, H., Kim, C.J., Shinn, P., Bowser, L., Carninci, P.,

Dale, J.M., Hayschizaki, Y., Huan, V.W., Ishida, J., Jones, T.,

Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lin, J., Miranda, M.,

Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L.,

Sakurai, T., Satou, M., Seki, M., Southwick, A., Toriumi, M., Wong, C.,

Wu, H.C., Yamada, K., Yu, G., Yuan, S., Shinozaki, K., Davis, R.W.,

Theologis, A. and Ecker, J.R.

Arabidopsis ORF clones

Unpublished

2 (bases 1 to 1959)

Chen, R., Chen, H., Kim, C.J., Shinn, P., Bowser, L., Carninci, P.,

Dale, J.M., Hayschizaki, Y., Huan, V.W., Ishida, J., Jones, T.,

Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lin, J., Miranda, M.,

Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L.,

Sakurai, T., Satou, M., Seki, M., Southwick, A., Toriumi, M., Wong, C.,

Wu, H.C., Yamada, K., Yu, G., Yuan, S., Shinozaki, K., Davis, R.W.,

Theologis, A. and Ecker, J.R.

Direct Submission

Submitted (15-MAY-2003)

Salk Institute Genomic Analysis Laboratory

(SIGNAL), Plant Biology Laboratory, The Salk Institute for

Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,

USA

The discrepancy does not affect the protein sequence.

RIKEN Genomic Sciences Center (GSC) members carried out the

collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN

Arabidopsis Full-Length cDNA') : Seki, M., Narusaka, M., Ishida, J.,

Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,

Hayschizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGEC (SSP) Consortium members constructed and

sequenced the PUNI (ORF) clones using the RAFL cDNAs: Chen, R.,

Chen, H., Kim, C.J., Shinn, P., Bowser, L., Chan, M.M., Chang, C.M.,

Dale, J.M., Huan, V.W., Jones, T., Karlin-Neumann, G., Lam, B.,

Lee, J.M., Lin, J., Miranda, M., Nguyen, M., Onodera, C.S., Palm, C.J.,

Quach, H.L., Southwick, A., Tang, C.C., Toriumi, M., Wong, C., Wu, H.C.,

Yamada, K., Yu, G., Yuan, S., Davis, R.W., Theologis, A., and

Ecker, J.R.

Chen, R. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to

this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk)

contributed equally to this work as PIs.

Location/Qualifiers

1. 1959

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/db_xref="taxon:3702"

/chromosome="3"

/clone="U17522"

/ecotype="Columbia"

/note="This clone is in PUNI 51"

FEATURES

source

CDS

1. .1959

/note="unknown protein"

/codon_start=1

/product="At3g24550"

/protein_id="AAP37759.1"

/db_xref="GI:30725474"

/translation="MSTAPSGTTPSPSPSPSTNTTTPPAASPPPTTTSSPP

PSSTNSTSPSPSPPLPPLPPSPSGSLTPPLPQSPSPAPITPSPSPPTSPSP

PSNQSPNTSPSTPTSPNTSPSPSSDGLSTGVVVGIAIGGVAILVILILCL

LCKKRRRHDDAAAYVPPPPSPGKAGPYGGQQYQQQNASRSDNHVVTSLPP

PKPPSPRPKPPPPPPPPAFMSSSGSDYSLPVLPPSPGLVLFSGSTFTYELSLR

TNGFSEANLQGGGVYVHKGILPSGKEVAVKOLKAGSGQEREFQAEVETISRVHHR

HLVSLGYCMAGVQRLVAVFVNNLEFLHKGKRPMTMSTELKIALGSAKGLYL

HEBCNPKIHRDIKASNILIDFPEAKVADFGLAKIASDNTNTHVSTVMGTFGYLAPE

YAAGSKUTEKSDVSPGVLELLITGRPPVDANNVYVDDSLVDWARLPLNRSSEGF

EGLASRMGNEYDREEMARMVACAAACVHSARRRPRMSQIVRALEGNVSLDNEGM

RPGHSNVYSSYGSGTDYDTSQYNDMDIKFRKMALGTQEGYGTGTSNPTSDYGLYPSG

SSSEGGQATREMENGKIKKTGGYSGPSL"

misc_difference 1920

/note="compared to genome"

/replace="t"

ORIGIN

Alignment Scores:

Pred. No.: 1.23e-63 Length: 1959

Score: 2936.50 Matches: 574

Percent Similarity: 90.61% Conservative: 24

Best Local Similarity: 86.97% Mismatches: 41

Query Match: 85.04% Indels: 21

DB: 8 Gaps: 12

US-10-086-464-2 (1-647) x BT008400 (1-1959)

Qy 1 MetSerAlaProSerProGlyThr-----GlySerProSerProSer 17

Db 1 ATGTCCACAGCGCGTCTCCAGGCACTACTCCATCCATCTCCACCGTCTCTCCACCA 60

Qy 18 AsnSerThrThrThrProProAla-----SerAlaProProThrPro 35

Db 61 AACTCGAACACCACTCTCTCCAGCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 120

Qy 36 SerSerProProPro-----ProSerThrIleProThrSerProProSerSerArgSer 54

Db 121 TCTTCT 180

Qy 55 ThrProSerAlaProProProProProProProProProProProProProProPro 74

Db 181 CCT 228

Qy 75 LeuProGlnProSerProProAlaProThrThrProGlySerProProAlaProValThr 94

Db 229 CTTCTCTCAACT 285

Qy 95 ProProThr---ArgAsnProProProSerValProGlyProProSerAsnProSerArg 113

Db 286 CCTCAAAACCTCGAAGCCCTCCATCTCTCAACCAAGGACCAACCAACACTCCCTCA 342

Qy 114 GluGlySerProArgProProProSerSer-----ProSerProProSerProSerSer 131

Db 343 ---GGATCTACTCTAGAACTCCATCAACACTAAACCGTCGCGCGCTCTGATCTCTCC 399

Qy 132 AspGlyLeuSerThrGlyValValValIleAlaIleGlyGlyValAlaLeuVal 151

Db 400 GATGGATGTCTACCGAGTGTGGTAGGAATCCCATTTGGAGGAGTCGCTATTCTTGT 459

Qy 152 IleValThrIleuIleCysLeuLeuCysLysLysLysArgArg-----AspGluGlu 169

Db 460 ATACTGACTCTGATTTGTTCTTCTCTGTAAAGAAACGAAGAACACACGACGATGAA 519

Qy 170 AspAlaTyrTyrValProProProProProProProProProProProProProPro 189

Db 520 GCTGCTTACTATGTTCT 579

```
Qy 190 GlyGlnGlnGlnTrpArgGlnAlaAsnAlaThrProProSerAsp---HisValVal 208
Db 580 GGTCAACAACAGTATTGGCAACAACAAACCGGTACCGCGCGTCAGATAATCATGTAGTG 639
Qy 209 ThrSerLeuProProProProGlyAlaProSerProProArgGlnProProProProPro 228
Db 640 ACGTCATTTG---CCACCACCTTAAGCCTCCATCTCCACCACCAACCACTCTCCGCCACCT 696
Qy 229 ProProPro---PheMetSerSerSerGlyGlySerAspTyrSerAspArgProValLeu 247
Db 697 CCACACACAGCATTCATGAGTAGCAGTGGTGGTCTGACTATTCGGATCTCCGGTCTTT 756
Qy 248 ProProProSerProGlyLeuValLeuGlyPheSerLysSerThrPheThrTyrGluGlu 267
Db 757 CCTCCACCATCTCCAGGCGTTGTGTAGGCTTTTCTAAAGACACTTTTCACTTATGAGGAG 816
Qy 268 LeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyGlyPheGly 287
Db 817 TTGTCGAGAGCTACTAATATGGCTTCTCTGAGGCTAAATTTGTTAGGACAGGAGGGTTTGGT 876
Qy 288 TyrValHisGlyGlyValLeuProSerGlyLysGluValAlaValLysGlnLeuValVal 307
Db 877 TATGTGCTAAGATATATGCTTAGTGGGAAGAGTTGCTGTGAACAGTGTGAAGCT 936
Qy 308 GlySerGlyGlnGlyGluArgGluPheGlnAlaGluValGluLeuLeuSerArgValHis 327
Db 937 GGTAGTGGTCAGGAGAGAGAGAGTTCAGGCTGAGGTTGAGATCATTAGCAGAGTTTCT 996
Qy 328 HisArgHisLeuValSerLeuValGlyTyrCysIleAlaGlyAlaLysArgLeuLeuVal 347
Db 997 CACAGGATTTGGTTTCTTATTGGTTATTGTTATGTCGCCGCTGTTCAAGATTAATTCTT 1056
Qy 348 TyrGluPheValProAsnAsnAsnLeuGluLeuHisGlyGlyGlyArgProThr 367
Db 1057 TATGAGTTTGTTCACAAACAACATCTTGAGTTTCACTCTCATGTGAAGAGCGCTACG 1116
Qy 368 MetGluTrpSerThrArgLeuLysIleAlaLeuGlySerAlaLysGlyLeuSerTyrLeu 387
Db 1117 ATGGAATGGAGTACTAGATTGAAGATTGCTCTTGATCTGCTAAAGGACITTCATATCTT 1176
Qy 388 HisGluAspCysAsnProLysIleIleHisArgAspIleLysAlaSerAsnIleLeuIle 407
Db 1177 CATGAAGATTGCATCCGAAATCATTCACCGTGATATTAAAGCGCTCAAAACATATTGATT 1236
Qy 408 AspPheLysPheGluAlaLysValAlaAspPheGlyLeuAlaLysIleAlaSerAspThr 427
Db 1237 GATTTCAAAATTTGAAGTAAAGTTGCTGACTTTGGTCTTGCCAGATGCTTCTTGATACA 1296
Qy 428 AsnThrHisValSerThrArgValMetGlyThrPheGlyTyrLeuAlaProGluTyrAla 447
Db 1297 AACACTCATGATCTACACGCGTGATGGAAACCTTTGGGTATTGGTCCGGAATATGCT 1356
Qy 448 AlaSerGlyLysLeuThrGluLysSerAspValPheSerPheGlyValValLeuLeuGlu 467
Db 1357 GCAAGTGGAAAGCTCACAGAAAAGCTGACGTTTCTCAITTTGGCGTTGATCTTTTGAA 1416
Qy 468 LeuIleThrGlyArgArgProValAspAlaAsnValTyrValAspAspSerLeuVal 487
Db 1417 CTTATTACTGGAGGCGCCCTGTTGATGCGCAACAATGCTCATGATGATGACAGCTTAGTT 1476
Qy 488 AspTriaAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGluGlyLeuAla 507
Db 1477 GACTGGGCACGACCATTTGCTTTACCGAGCATCTGAGGAAGAGATTTTGGGGTTGGCT 1536
Qy 508 AspAlaLysMetAsnAsnGlyTyrAspArgGluGluMetAlaArgMetValAlaCysAla 527
Db 1537 GATTCAAAGATGGGTAATAGTATGACAGAGAGAGATGCTCGCATGGTTGCTTGGCT 1596
Qy 528 AlaAlaCysValArgHisSerAlaArgArgProArgMetSerGlnIleValArgAla 547
Db 1597 GCGGCTTGTGTTGCCATTGAGCTCGCCGAGAGCTCGCATGAGCCAGATAGTACGGGCG 1656
Qy 548 LeuGluGlyAsnValSerLeuSerAspLeuAsnGluGlyMetArgProGlyGlnSerAsn 567
```

```
Db 1657 TTAGACGAAATGATCGCTGTCTGATCTTAACGACGAGATGACACCGGTCACAGCAAC 1716
Qy 568 ValTyrSerSerTyrGlyGlySerThrAspTyrAspSerSerGlnTyrAsnGluAspMet 587
Db 1717 GTATACAGCTCATATGAGGAAGACACAGACTATGACAGGACCAATACACAGCAGCATG 1776
Qy 588 LysLysPheArgLysMetAlaLeuGlyThrGlnGluTyrAsnAlaThrGlyGluTyrSer 607
Db 1777 ATAAAGTTTAGGAATATGGCTCTTGGAACTCAAGNATACGGCACAAACGGGAGTACAGT 1836
Qy 608 AsnProThrSerAspTyrGlyLeuTyrProSerGlySerSerGlyGluGlnThrThr 627
Db 1837 AATCCAAACCACTGACTACGAGCTGTACCGCTCTGTTCAAGCAGTGAAGGTCAAGCCACA 1896
Qy 628 ArgGluMetGluMetGlyLysIleLysArgThrGlyGlnGlyTyrSerGlyProSerLeu 647
Db 1897 CGAGNAATGAGATGGGAAAGATCAAGAAAACCGGTCAAGGTTATAGTGGACCTCTCTT 1956

RESULT 7
LOCUS AV128792 2098 bp mRNA linear PLN 07-AUG-2002
DEFINITION Arabidopsis thaliana protein kinase-like protein mRNA, complete cds.
ACCESSION AV128792
VERSION AV128792.1 GI:22136227
KEYWORDS FLJ CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
REFERENCE 1. (bases 1 to 2098)
AUTHORS Tripp,M., Southwick,A., Karlin-Neumann,G., Nguyen,M., Miranda,M.,
Palm,C.J., Bowser,L., Jones,T., Banh,J., Carninci,P., Chen,H.,
Chen,K., Chung,M.K., Hayashizaki,Y., Ishida,J., Kamiya,A.,
Kawai,J., Kim,C., Lin,J., Liu,S.X., Narusaka,M., Pham,P.K.,
Sakano,H., Sakurai,T., Satou,M., Seki,M., Shinn,P., Yamada,K.,
Shinozaki,K., Ecker,J., Theologis,A. and Davis,R.W.
Direct Submission
JOURNAL Submitted (01-JUL-2002) DNA Sequencing and Technology Center,
Stanford University, 855 California Avenue, Palo Alto, CA 94304,
USA
COMMENT e-mail for correspondence: arab@sequence.stanford.edu

The RIKEN Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAFL cDNAs (RAFL cDNA: 'RIKEN
Arabidopsis Full-length cDNA'): Seki,M., Narusaka,M., Ishida,J.,
Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PDEC (SSP) Consortium members constructed and
sequenced the PUNI (ORF) clones using the RAFL cDNAs: Tripp,M.,
Nguyen,M., Southwick,A., Karlin-Neumann,G., Lam,B., Miranda,M.,
Palm,C.J., Bowser,L., Jones,T., Banh,J., Chen,H., Chen,K.,
Chung,M.K., Kim,C., Lin,J., Liu,S.X., Pham,P.K., Sakano,H.,
Shinn,P., Yamada,K., Ecker,J., Theologis,A. and Davis,R.W.

Tripp,M. (SSP/Stanford) and Seki,M. (RIKEN GSC) contributed equally
to this work. Shinozaki,K. (RIKEN GSC) and Davis,R.W.
(SSP/Stanford) contributed equally to this work as PIs.

FEATURES
source
1. .2098
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="U16568"
/ecotype="Columbia"
/note="This clone is in pENTR/SD-dTopo This is a cloned
PCR product using RIKEN clone RAFL09-06-A17 (AY059901) as
a template"
1. .1959
/codon_start=1

CDS
```



```
/product="protein kinase-like protein"
/protein_id="AAM51192.1"
/db_xref="GI:22136228"
/translation="MSTAPSGFTSPSPSPPTNTTTPPPAASSPPPTTTPSSPP
PSPNTGTPSPSPSPPLPSPSPGSLTPLPQSPSPAPITPSPSPPTPSPNRPSP
PSNQGTPNTSPSGTPTPSTKPSPSDSGLSTGVVGIAGVAILVILTLICL
LCKKRRRHDDAAYVVPVPPSPGKAGPYGQQQYMQQNASRSPDNHVTSLLP
PKPSPRKPPPPPPAFMSSGGSYSDLPVLPSPGLVLGFSKSTTYBELSRA
TPGSEANLVGGQGFVYHKGIVLPSGKEVAVKQKAGSGQGERFOAEVILSRVHR
HLVSLIGYCMAGVORLLVYEFVFNNNLEFHLHGKRTMEMSTRLKALGSAKLSYL
HEDCNPKI IHRDIDKASNILIDFPEAKVADFLAKIASDNTNTHVSTVMGTGFLAIE
YAASGKTKEDSDVFSFGVLELITGRPPVDANNVYDDSLTWDWARPLLNARSEEGP
EGLADSKMGNEYDREEMRWACAAACVRHSARERPRMSQIVRALEGNLSLDNEGM
RPHSNVYSSITGGSDTDTQINDDMIKFRKRWALGTQBYGTGTEYSNPTSDIYLPSG
SSSEGOATREMENMKIKKTGGYSGPSL"
```

ORIGIN

Alignment Scores:

```
Pred. No.: 1,29e-63 Length: 2098
Score: 2936.50 Matches: 574
Percent Similarity: 90.61% Conservative: 24
Best Local Similarity: 86.97% Mismatches: 41
Query Match: 85.04% Indels: 21
DB: 8 Gaps: 12
```

US-10-086-464-2 (1-647) x AV128792 (1-2098)

```
Qy 1 MetSerSerAlaProSerProGlyThr-----GlySerProProSerProSer 17
Db 1 ATGTCACAGCGCGTCTCCAGGCACTACTCCATCACCATCTCCACCGCTCTCCACCA 60
Qy 18 AsnSerThrThrThrProProProAla-----SerAlaProProThrThrPro 35
Db 61 AACTCGACACACCACTCTCTCCAGCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 120
Qy 36 SerSerProProPro---ProSerThrThrThrThrThrThrThrThrThrThr 54
Db 121 TCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180
Qy 55 ThrProSerAlaProProProSerProProThrProThrProGlySerProPro 74
Db 181 CCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 228
Qy 75 LeuProGlnProSerProProAlaProThrProGlySerProProAlaProValThr 94
Db 229 CTTCCTCAACCTTCCCGCTCCGCTCCATCACTCTCTCTCTCTCTCTCTCTCTCTCT 285
Qy 95 ProProThr---ArgAsnProProSerProValProGlyProProSerAsnProSer 113
Db 286 CCTCAAAACCTCGAAGCCCTCCATCTCTTAACCAAGGACCAACCACTCTCTCTCA 342
Qy 114 GluGlyGlySerProArgProProSerSer-----ProSerProProSerProSer 131
Db 343 ---GGATCTACTCTAGAACTCCATCAACACTAAACCGTCGCGCGCTCTGATCTCTCC 399
Qy 132 AspGlyLeuSerThrGlyValValValGlylleAlaIleGlyGlyValAlaLeuVal 151
Db 400 GATGGATTGTCTACCGGAGTTGTGTAGGAATCGCCATTGGAGGAGTCGCTATTCTTGT 459
Qy 152 IleValThrLeuIleCysLeuLeuCysIleValValValValValValValValVal 169
Db 460 ATACTGACTCTGATTGTCTCTCTCTGTAAGAAGAAACGAAGAAGACACGACGATGAA 519
Qy 170 AspAlaTyrThrValProProProProProGlyProGlyProGlyProGlyProGly 189
Db 520 GCTGCTTACTATGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 579
Qy 190 GlyGlnGlnGlnThrArgGlnGlnAsnAlaThrProProSerAsp---HisValVal 208
Db 580 GGTCAACAACAGTATTGGCAACAAACAAACCGCTCAACCGCGCTCAGATTAATCATGT 639
Qy 209 ThrSerLeuProProProProProProProProProProProProProProProPro 228
```

```
Db 640 ACGTCAATGG---CCACCAACCTAAGCCTCCATCTCCACCAAGAAAACCTCTCCGCCACCT 696
Qy 229 ProProPro---PheMetSerSerSerGlyGlySerAspTyrSerAspArgProValLeu 247
Db 697 CCACCAACCAATTCATGAGTAGCAGTGGTGGTCTGACTATTTCGGATCTTCCGGTCTT 756
Qy 248 ProProProSerProGlyLeuValLeuGlyPheSerLysSerThrPheThrTyrGluGlu 267
Db 757 CCTCCACCATCTCCAGGCGCTTGTGTAGGCTTTCTTAAAGCACTTTTCACTTAITAGAGAG 816
Qy 268 LeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyPheGly 287
Db 817 TTGTCGAGAGCTACTAATGCTTCTCTGAGGCTAAATTTGTTAGCAAGGAGGTTTGGT 876
Qy 288 TyrValHisLysGlyValLeuProSerGlyLysGluValAlaValLysGlnLeuLysVal 307
Db 877 TATGTGCATAAAGTATATTGCTTAGTGGGAAAGAGTGTCTGTGAAACAGTTGAAAGCT 936
Qy 308 GlySerGlyGlnGlyGluArgGluPheGlnAlaGluValGluIleIleSerArgValHis 327
Db 937 GGTAGTGGTCAGGAGAGAGAGAGTTTCAGGCTGAGGTTGAGATCATTTAGCAGAGTTTAT 996
Qy 328 HisArgHisLeuValSerLeuValGlyTyrCysIleAlaGlyAlaLysArgLeuVal 347
Db 997 CACAGGCATTTGGTTTCTTATTGGTTATTGATGCGCGGTGTTCAAAGATTACTTGTGTC 1056
Qy 348 TyrGluPheValProAsnAsnAsnLeuGluLeuHisGlyGlyGlyGlyArgProThr 367
Db 1057 TATGAGTTGTTTCCAAACAACTTCTGAGTTTCACTCTCCATGTTAAAGGAGCGCTACG 1116
Qy 368 MetGluTrpSerThrArgLeuLysIleAlaLeuGlySerAlaLysGlyLeuSerTyrLeu 387
Db 1117 ATGGAATGGAGTACTAGATTGAAGATTGCTCTTGGATCTGCTAAAGGACITTCATATCTT 1176
Qy 388 HisGluAspCysAsnProLysIleIleHisArgAspIleLysAlaSerAsnIleLeu 407
Db 1177 CATGAAGATTGCAATCCGAAATCATTCACCGTGATTAAGGCGTCAACATATTGATT 1236
Qy 408 AspPheLysPheGluAlaLysValAlaAspPheGlyLeuAlaLysIleAlaSerAspThr 427
Db 1237 GATTTCAAATTTGAAGTAAAGTTGCTGACTTTTGGTCTTGCCAAAGATTGCTTCTGATACA 1296
Qy 428 AsnThrHisValSerThrArgValMetGlyThrPheGlyTyrLeuAlaProGluTyrAla 447
Db 1297 AACACTCATGTATCTCACCGGTGATGGAAACCTTTGGGTATTGGTTCGGAATATGCT 1356
Qy 448 AlaSerGlyLysLeuThrGluLysSerAspValPheSerPheGlyValValLeuLeuGlu 467
Db 1357 GCAAGTCGAAAGCTCACAGAAAAGCTCTGACGTTTCTCATTTGGCGTTGTACTTTTGGAA 1416
Qy 468 LeuIleThrGlyArgArgProValAspAlaAsnAsnValTyrValAspAspSerLeuVal 487
Db 1417 CTTATTATCTGGAGGCGCTCTGTTGATCGCAACAATGCTCTATGTAGATGACAGCTTAGTT 1476
Qy 488 AspTrpAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGlyLeuAla 507
Db 1477 GACTGGGACGACCATTTGCTTTAACCGAGCATCTGAGGAGAGAGATTTTGAGGGTTTGCT 1536
Qy 508 AspAlaLysMetAsnAsnGlyTyrAspArgGluMetAlaArgMetValAlaCysAla 527
Db 1537 GATTCAAAGATGGGTATGAGTATGACAGAGAGAGATGGCTCGCATGGTTGCTTGGCT 1596
Qy 528 AlaAlaCysValArgHisSerAlaArgArgProArgMetSerGlnIleValArgAla 547
Db 1597 GCGGCTTGTGTCCCATTCAGCTCGCGCAGACCTCGCATGAGCCAGATAGTAGCGGCG 1656
Qy 548 LeuGluGlyAsnValSerLeuSerAspLeuAsnGluGlyMetArgProGlyGlnSerAsn 567
Db 1657 TTAGAAGAAATGATCGCTGCTGATCTTAACGAAGGATGAGACCGGTCACAGCAAC 1716
Qy 568 ValTyrSerSerTyrGlySerThrAspTyrAspSerSerGlnTyrAsnGluAspMet 587
Db 1717 GTATACAGCTCATATGAGGAAGACACAGACTATGACACGAGCAATATCAACAGCACGACATG 1776
```


Db 556 GCTGTTACTATGTTCTCTCTCTCCATCTGGTCCCAAGCTGGAGGACCTTACGGT 615
 Qy 190 GlyGlnGlnGlnTrpArgGlnGlnAsnAlaThrProProSerAsp---HisValVal 208
 Db 616 GGTCAACAACAGTATTGGCAACAACAAACCGCTCAGCCGCTCAGATAATCATGTAGTG 675
 Qy 209 ThrSerLeuProProProProLysAlaProSerProProArgGlnProProProPro 228
 Db 676 ACGTCAATG---CCACCACCTTAAGCCTCCATCTCCACCACCAAAACCTCTCCGCCACCT 732
 Qy 229 ProProPro---PheMetSerSerGlyGlySerAspTyrSerAspArgProValLeu 247
 Db 733 CCACCACGAGCATTCATGAGTATGAGTGGTGTCTGACTATTCCGATCTTCCGGTCTT 792
 Qy 248 ProProProSerProGlyLeuValLeuGlyPheSerLysSerThrPheThrTyrGluGlu 267
 Db 793 CCTCCACCATTCTCCAGGGCTTGTGTAGGCTTTTCTAAAGCACTTTTCAATATGAGGAG 852
 Qy 268 LeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyPheGly 287
 Db 853 TTGTCCGAGAGTACTAATGGCTTCTCTGAGGCTAATTTGTAGGACAGGAGGGTGTGGT 912
 Qy 288 TyrValHisLysGlyValLeuProSerGlyLysGluValAlaValLysGlnLeuLysVal 307
 Db 913 TATGTGCATAAAGGTATATTGCTAGTGGGAAAGAGTTGCTGTGAAACAGTTGAAGCT 972
 Qy 308 GlySerGlyGlnGlyGluArgGluPheGlnAlaGluValGluLulleSerArgValHis 327
 Db 973 GGTAGTGTGTCAGGAGAGAGAGAGTTTCAGGCTGAGGTTGAGATCATTTAGCAGAGTTTCA 1032
 Qy 328 HisArgHisLeuValSerLeuValGlyTyrCysLleAlaGlyAlaLysArgLeuVal 347
 Db 1033 CACAGGCATTTGGTGTCTCTTTATGGTATTGTATGGCCGGTGTCAAGAGTATTACTTGT 1092
 Qy 348 TyrGluPheValProAsnAsnLeuGluLeuHisLeuHisGlyGluGlyArgProThr 367
 Db 1093 TATGATTTGTTCNAAACACAATCTTGAGTTTACCTTCATGTGTAGGGACGGCTACG 1152
 Qy 368 MetGluTrpSerThrArgLeuLysLleAlaLeuGlySerAlaLysGlyLeuSerTyrLeu 387
 Db 1153 ATGGAATGGAGTACTAGATTCAAGATTGCTCTTGATCTGTCTAAAGGACTTTCATATCTT 1212
 Qy 388 HisGluAspCysAsnProLysLleIleHisArgAspLleLysAlaSerAsnLleIle 407
 Db 1213 CATGAAGATTGCAATCCGAAATCATTCACCGTGATATTAAAGGGCTCAAAATATTGATT 1272
 Qy 408 AspPheLysPheGluAlaLysValAlaAspPheGlyLeuAlaLysLleAlaSerAspThr 427
 Db 1273 GATTTCAATTTGAAGCTAAGTTGCTGACTTTGGTCTTGGCAAGATTGCTTCTGATACA 1332
 Qy 428 AsnThrHisValSerThrArgValMetGlyThrPheGlyTyrLeuAlaProGluTyrAla 447
 Db 1333 AACACTCATGTATCTACACGGCTGATGGAAACCTTTGGGTATTGCTCCGGAATATGCT 1392
 Qy 448 AlaSerGlyLysLeuThrGluLysSerAspValPheSerPheGlyValValLeuLeuGlu 467
 Db 1393 GCAAGTGGAAAGCTCACAGAAAGTCTGACGTTTCTCATTTGGCGTTGTACTTTTGAA 1452
 Qy 468 LeuLleThrGlyArgArgProValAspAlaAsnAsnValTyrValAspSerLeuVal 487
 Db 1453 CTTATTACTGGGAGGGCCCTGTTGATCGCAACATGTCTATGTAGATGACAGCTTAGTT 1512
 Qy 488 AspTrpAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGluGlyLeuAla 507
 Db 1513 GACTGGGCACGACCATTCCTTAACCGAGCATCTCAGGAAGAGAGATTTTGGGGTGTGGCT 1572
 Qy 508 AspAlaLysMetAsnAsnGlyTyrAspArgGluGluMetAlaArgMetValAlaCysAla 527
 Db 1573 GATTCAAGATGGGTAATGAGTATGACAGAGAGAGATGGCTCGCATGTTGGTTCGGCT 1632
 Qy 528 AlaAlaCysValArgHisSerAlaArgArgArgProArgMetSerGlnIleValArgAla 547
 Db 1633 CGGGCTTGTTCGCATTCAGCTCGCCGACAGCTCGCATGAGCCAGATAGTACGGGCG 1692

Qy 548 LeuGluGlyAsnValSerLeuSerLeuSerAspLeuAsnGluGlyMetArgProGlyGlnSerAsn 567
 Db 1693 TTAGAAGGAATGTATCGCTGTCTGATCTTAAACAGAGGGATGAGACCGGGTCAAGCAAC 1752
 Qy 568 ValTyrSerSerTyrGlyGlySerThrAspTyrAspSerSerGlnTyrAsnGluAspMet 587
 Db 1753 GTATACAGCTCATATGAGGAGGACACAGACTATGACACGAGCCCAATCAACAGCAGCATG 1812
 Qy 588 LysLysPheArgLysMetAlaLeuGlyThrGlnGluTyrAsnAlaThrGlyGluTyrSer 607
 Db 1813 ATAAAGTTTAGGAAATGGCTCTTTGGAACTCAAGATACGGCACAAACCGGCGAGTACAGT 1872
 Qy 608 AsnProThrSerAspTyrGlyLeuTyrProSerGlySerSerSerGluGlyGlnThrThr 627
 Db 1873 AATCAACCACTGACTACGAGCTGTACCGCTCTGGTTCAAGCAGTGAAGGTCAAGCCACA 1932
 Qy 628 ArgGluMetGluMetGlyLysLleLysArgThrGlyGlnGlyTyrSerGlyProSerLeu 647
 Db 1933 CGAGAAATGGAGATGGGAAGATTAAAGAAACCGGTCAAGTTATAGTGGACCTCTCTT 1992
 RESULT 9
 AY059901 2188 bp mRNA linear PLN 04-NOV-2001
 LOCUS Arabidopsis thaliana protein kinase-like protein (MOB24.13) mRNA,
 DEFINITION complete cds.
 ACCESSION AY059901
 VERSION AY059901.1 GI:16649062
 KEYWORDS FLI CDNA.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopses.
 1 (bases 1 to 2188)
 AUTHORS Nguyen, M., Karlin-Neumann, G., Southwick, A., Lam, B., Miranda, M.,
 Palm, C.J., Bowser, L., Jones, T., Banh, J., Carninci, P., Chen, H.,
 Cheuk, R., Chung, M.K., Hayashizaki, Y., Ishida, J., Kamiya, A.,
 Kawai, J., Kim, C., Lin, J., Liu, S.X., Narusaka, M., Pham, P.K.,
 Sakano, H., Sakurai, T., Satou, M., Seki, M., Shinn, P., Yamada, K.,
 Shinozaki, K., Ecker, J., Theologis, A. and Davis, R.W.
 DIRECT SUBMISSION
 TITLE Submitted (22-OCT-2001) DNA Sequencing and Technology Center,
 JOURNAL Stanford University, 855 California Avenue, Palo Alto, CA 94304,
 COMMENT e-mail for correspondence: arab@sequence.stanford.edu
 RIKEN Genomic Sciences Center (GSC) members carried out the
 collection and clustering of RAPL cDNAs (RAPL cDNA : 'RIKEN
 Arabidopsis Full-length cDNA'): Seki, M., Narusaka, M., Ishida, J.,
 Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,
 Hayashizaki, Y. and Shinozaki, K.
 The Salk, Stanford, PGSC (SSP) Consortium members carried out the
 sequencing and annotation of the RAPL cDNAs: Nguyen, M.,
 Southwick, A., Karlin-Neumann, G., Lam, B., Miranda, M., Palm, C.J.,
 Bowser, L., Jones, T., Banh, J., Chen, H., Cheuk, R., Chung, M.K.,
 Kim, C., Lin, J., Liu, S.X., Pham, P.K., Sakano, H., Shinn, P.,
 Yamada, K., Ecker, J., Theologis, A. and Davis, R.W.
 Nguyen, M. (SSP/Stanford) and Seki, M. (RIKEN GSC) contributed
 equally to this work. Shinozaki, K. (RIKEN GSC) and Davis, R.W.
 (SSP/Stanford) contributed equally to this work as PIs.
 FEATURES
 source
 1. 2188
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /chromosome="III"
 /clone="RAPL09-06-A17"
 /note="This clone is in pBluescript"
 1. 2188
 /gene="MOB24.13"
 gene

CDS

```

32..1990
/gene="MOB24.13"
/codon_start=1
/product="protein kinase-like protein"
/db_xref="GI:16649063"
/translacion="MSTAPSPGTPSPSPPTNTTTPPPAASPPPTTPSPSP
PSPNTSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP
PSNQGPNTPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP
LCKKRRRHDDDEAAYVPPPPSPGPKAGPYGQQQWQQQNASRSDHNVTSLLP
PKPSPKPPPPPPPPAFMSSGSDYSDLPVLPSPSPGLVLGFSKSTFTYELSRA
TNGSEANLGGQGVYHKGILPSGKEVAVKQKAGSGQGERFOAEVIEIKRVHR
HLVSLIGYAGVORLLVYEFVNNNLEFHLHGKRTMESTRLKIALGSAKGLSYL
HDCNPKIIRHDIKASNLIDPFEKAVADFGLAKIASDTNTHVSTVMGTGFLADE
YAASGKLTKSDVPSGVLLLELTGRPPVDANNVDDSLTVDWARPLANRASEGDF
EGLADSKMGNEYDEEMARVACAAACVRSARRPRMSQIVRALEGNVSLDLNEM
RPHSHSVYSSGGSDYDSTQYNDMDIKFRKMLGTOEYGTGTGEYSNPTSDYGLYPSG
SSSEGOATREMEMGKIKKTQGYSGPSL"

```

ORIGIN

Alignment Scores:

Pred. No.:	1.33e-63	Length:	2188
Score:	2936.50	Matches:	574
Percent Similarity:	90.61%	Conservative:	24
Best Local Similarity:	86.97%	Mismatches:	41
Query Match:	85.04%	Indels:	21
DB:	8	Gaps:	12

US-10-086-464-2 (1-647) x AY059901 (1-2188)

```

Qy 1 MetSerSerAlaProSerProGlyThr-----GlySerProProSerProSer 17
Db 32 ATGTCACAGCGCGCTCTCCAGGCACTACTCCATCACCATCTCCACCGCTCTCCACCA 91
Qy 18 AsnSerThrThrThrProProAla-----SerAlaProProThrThrPro 35
Db 92 AACTCGACAACCACTCTCTCTCCAGCAGCCTCTCTCTCTCTCTCTCTCTCTCTCTCT 151
Qy 36 SerSerProProPro-----ProSerThrThrThrThrThrThrThrThrThrThr 54
Db 152 TCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 211
Qy 55 ThrProSerAlaProProProSerProProThrProSerProGlySerProProPro 74
Db 212 CCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 259
Qy 75 LeuProGlnProSerProAlaProThrThrProGlySerProProAlaProValThr 94
Db 260 CTTCCTCAACCTTCCCGCTCCCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 316
Qy 95 ProProThr---ArgAsnProProSerValProGlyProProSerAsnProSerArg 113
Db 317 CCCTCAAAACCTCGAAGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 373
Qy 114 GluGlyGlySerProArgProProSerSer-----ProSerProProSerProSer 131
Db 374 ---GGATCTACTCTAGACTCTCACTCAAACTAAACCGTCCCGCGCTCTGATCTCTCC 430
Qy 132 AspGlyLeuSerThrGlyValValValGlyLeuAlaLeuGlyValAlaLeuVal 151
Db 431 GATGGATTGTCTACCGGAGTTGTGTAGGATCCCATCTCTCTCTCTCTCTCTCTCTCT 490
Qy 152 IleValThrLeuIleCysLeuLeuCysLeuValValValValValValValValValVal 169
Db 491 ATACTGACTCTGATTTGTCTCTCTGTAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 550
Qy 170 AspAlaValThrValProProProProProProProProProProProProProPro 189
Db 551 GCTGCTTACTATGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 610
Qy 190 GlyGlnGlnGlnIntrpArgGlnGlnAsnAlaThrProProSerAsp---HisValVal 208
Db 611 GGTCAACAACAGTATTGGCAACAACAAACCGCGTCACGCGCGTCAGATATCATGTAGTG 670

```

```

Qy 209 ThrSerLeuProProProProLysAlaProSerProProArgGlnProProProPro 228
Db 671 ACGTCAITG---CCACCACCTTAGCCTCCATCTCCACCAAGAAACCTCTCTCCGCCACCT 727
Qy 229 ProProPro---PheMetSerSerSerGlyGlySerAspTyrSerAspArgProVal 247
Db 728 CCACCACAGCATTTCATGAGTAGCAGTGGTGTCTGACTATTCCGATCTTCCGGTCTT 787
Qy 248 ProProProSerProGlyLeuValLeuGlyPheSerLysSerThrPheThrThrGluGlu 267
Db 788 CTTCCACCATCTCCAGGCTTGTGTAGGCTTTTCTAAAGACATTTTACATTTATGAGGAG 847
Qy 268 LeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyPheGly 287
Db 848 TTGTCGAGAGCTACTAATGCTTCTCTGAGCTAATTTGTAGGACCAAGGAGGTTTGT 907
Qy 288 TyrValHisLysGlyValLeuProSerGlyLysGluValAlaValLysGlnLeuLysVal 307
Db 908 TATGTGCATAAAGGTATATTGCTTAGTGGGAAAGAGTGTCTGTGAACAGTTCGAAGCT 967
Qy 308 GlySerGlyGlnGlyGluArgGluPheGlnAlaGluValGluIleIleSerArgValHis 327
Db 968 GGTAGTGTCTCAGGAGAGAGAGAGTTCAGGCTGAGGTAGATCATTTAGCAGAGTTTCT 1027
Qy 328 HisArgHisLeuValSerLeuValGlyTyrCysIleAlaGlyAlaLysArgLeuVal 347
Db 1028 CACAGGCAATTTGTTCTTATTGTTATTTGATGCGCGGTGTCTCAAGATTTACTTGT 1087
Qy 348 TyrGluPheValProAsnAsnAsnLeuGluLeuHisGlyGlyGlyGlyGlyGlyGly 367
Db 1088 TATGAGTTGTTTCCAAACAACATCTTGTAGTTTCCCTCCATGTTAAGGAGCGCTACG 1147
Qy 368 MetGluTrpSerThrArgLeuLysIleAlaLeuGlySerAlaLysGlyLeuSerTyrLeu 387
Db 1148 ATGGAATGAGTACTAGATTGAAGATTGCTCTTGGATCTGCTAAAGGACTTTTATATCT 1207
Qy 388 HisGluAspCysAsnProLysIleIleHisArgAspIleLysAlaSerAsnIleLeu 407
Db 1208 CATGAAGATTGCAATCCGAAATCATTTACCGTGATATTAAAGGCTCAACATATTGATT 1267
Qy 408 AspPheLysPheGluAlaLysValAlaAspPheGlyLeuAlaLysIleAlaSerAspThr 427
Db 1268 GATTTCAAATTTGAAGCTAAGTTGCTGACTTTTGGTCTTGCCTGCTTCTCTGATACA 1327
Qy 428 AsnThrHisValSerThrArgValMetGlyThrPheGlyTyrLeuAlaProGluTyrAla 447
Db 1328 AACACTCATGTATCTACACGCTGATGGAACCTTTGGGTATTTGGGTCCGGAATATGCT 1387
Qy 448 AlaSerGlyLysLeuThrGluLysSerAspValPheSerPheGlyValValLeuLeuGlu 467
Db 1388 GCAAGTGAAGAGCTCACAGAAAGCTGAGCTTTTCTCATTTGGCGTTGTACTTTGGAA 1447
Qy 468 LeuIleThrGlyArgArgProValAspAlaAsnAsnValTyrValAspAspSerLeuVal 487
Db 1448 CTTATTACTGGAGGCGCCCTGTTGATGCGCAACATGCTATGTATAGATGACAGCTTAGTT 1507
Qy 488 AspTrpAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGluGlyLeuAla 507
Db 1508 GACTGGGCACGACCATTTGCTTAACCCGACATCTGAGGAAGAGATTTTGGAGGTTTGGCT 1567
Qy 508 AspAlaLysMetAsnAsnGlyTyrAspArgGluGluMetAlaArgMetValAlaCysAla 527
Db 1568 GATTCAAGATGGGTAAATAGTATGACAGAGAGAGATGCTCGCATGTTGTTGCTTGCCT 1627
Qy 528 AlaAlaCysValArgHisSerAlaArgArgProArgProArgMetSerGlnIleValArgAla 547
Db 1628 GCGGCTTGTGTTGCTTCCATTCAGCTCCGCGCAGACCTCGCATGAGCCAGATAGTAGCGG 1687
Qy 548 LeuGluGlyAsnValSerLeuSerAspLeuAsnGluGlyMetArgProGlyGlnSerAsn 567
Db 1688 TTGAAGGAAATGTATGCTGTCTGATCTTAAACGAAGAGGATGAGACCGGGTCACAGCAAC 1747

```

```

Qy 568 ValTyrSerSerTyrGlySerThrAspTyrAspSerSerGlnTyrAsnGluAspMet 587
Db 1748 GTATACAGCTCATATGGAGGAAGCACAGACTATGACAGAGCCCAATCAACGACGACATG 1807
Qy 588 LysIysPheArgLysMetAlaLeuGlyThrGlnGluTyrAsnAlaThrGlyGluTyrSer 607
Db 1808 ATAAAGTTTAGAAAATGGCTCTTGGAACCTCAAGNATACGCACACCGCGGAGTACAGT 1867
Qy 608 AsnProThrSerAspTyrGlyLeuTyrProSerGlySerSerSerGluGlyGlnThrThr 627
Db 1868 AATCCAAACAGTACTACGAGCTACCGCTCTGGTTCAGACAGTGAAGGTCAAGCCACA 1927
Qy 628 ArgGluMetGluMetGlyLysAlaLeuLysArgThrGlyGlnGlyTyrSerGlyProSerLeu 647
Db 1928 CGAGAAATGGAGATGGGAAGATTAAAGAAACCGGTCAAGGTATTATAGTGGACCTCTCTT 1987

```

```

RESULT 10
AY093065
LOCUS AY093065 2190 bp mRNA linear PLN 21-APR-2002
DEFINITION Arabidopsis thaliana unknown protein (At3g24550) mRNA, complete cds.
ACCESSION AY093065
VERSION AY093065.1 GI:20260331
KEYWORDS FLI CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
REFERENCE
AUTHORS Southwick,A., Karlin-Neumann,G., Nguyen,M., Lam,B., Miranda,M., Palm,C.J., Bowser,L., Jones,T., Banh,J., Carninci,P., Chen,H., Cheuk,R., Chung,M.K., Hayashizaki,Y., Ishida,J., Kamiya,A., Kawai,J., Kim,C., Lin,J., Liu,S.X., Narusaka,M., Pham,P.K., Sakano,H., Sakurai,T., Satou,M., Seki,M., Shinn,P., Yamada,K., Shinozaki,K., Ecker,J., Theologis,A. and Davis,R.W.
DIRECT SUBMISSION
SUBMITTED (27-MAR-2002) DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA
COMMENT e-mail for correspondence: arab@sequence.stanford.edu

```

RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL CDNA : 'RIKEN Arabidopsis Full-length cDNA') : Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGSC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Southwick,A., Nguyen,M., Karlin-Neumann,G., Lam,B., Miranda,M., Palm,C.J., Bowser,L., Jones,T., Banh,J., Chen,H., Cheuk,R., Chung,M.K., Kim,C., Lin,J., Liu,S.X., Pham,P.K., Sakano,H., Shinn,P., Yamada,K., Ecker,J., Theologis,A. and Davis,R.W.

Southwick,A., (SSP/Stanford) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Davis,R.W. (SSP/Stanford) contributed equally to this work as PIs.

```

FEATURES
source
    1..2190
        /organism="Arabidopsis thaliana"
        /mol_type="mRNA"
        /db_xref="taxon:3702"
        /chromosome="3"
        /clone="RAFL07-12-E08 (R17522)"
        /ecotype="Columbia"
        /note="This clone is in pBluescript"
    1..2190
        /gene="At3g24550"
    30..1988
        /gene="At3g24550"
        /codon_start=1
        /product="unknown protein"

```

```

gene
    30..1988
        /gene="At3g24550"

```

```

CDS
    30..1988
        /gene="At3g24550"
        /codon_start=1
        /product="unknown protein"

```

```

/protein_id="AAAM13064.1"
/db_xref="GI:20260332"
/translators="MSTAPSPGTTTTPSPSPPTNSTTTTTPPAASPPPTTTTSSPP
PSSTNSTSPSPSPPLPSPSPGSLTLPSPSPAPITPSPSPPTTSPSPSP
PSNQGPNTPTSGTPTPTSPSPSPSDSLSTGVVGIAGVAILVILICL
LCKKRRRRHDDAAAYVPPPPSPKAGPGYQQYQQQNASRSDNHVTSLLP
PKGSPRPKPPPPPPPPSPSSGSDYSDLPVLPSPGLVLFKFSKFTYEBSRA
TNGFSEANLGGGFGVVKGIILPSGKEVAVKOLKAGSGGGEREFOAEVILSRVHR
HLVSLGYNAGVQRLLVYFVFNNNLEFHLHGKRTMWSLTKALASGKLSYL
HEDCNPKI IHRDKASNLIDLFPEAKVDFGLAKIASDINTHVWMTGTFGLAPE
YAAAGKLTEDKVSFSGVLELITGRPPDANNVYVDDSLVWARPLNARSEGEF
EGLASKMGNEYDREEMARVCAACACVRSARRPRMSQIVRALEGNVSLDINEM
RGHNSVYSSGYSTDYDTSQVNDMDMKFRKMALGTQYGTGTGEYSNPTSGLYPSG
SSSEGQATREMENGKIKKTCQYSGPSL"

```

ORIGIN

```

Alignment Scores:
Pred. No.: 1.33e-63 Length: 2190
Score: 2936.50 Matches: 574
Percent Similarity: 90.61% Conservative: 24
Best Local Similarity: 86.97% Mismatches: 41
Query Match: 85.04% Indels: 21
DB: 8 Gaps: 12

US-10-086-464-2 (1-647) x AY093065 (1-2190)

Qy 1 MetSerSerAlaProSerProGlyThr-----GlySerProProSerProProSer 17
Db 30 ATGTCCACAGCGCGTCTCCAGGCACACTACTCCATCCACCATCTCCACCGTCTCTCC 89
Qy 18 AsnSerThrThrThrProProAla-----SerAlaProProThrThrPro 35
Db 90 AACTCGAACCACCATCTCTCTCCAGAGCCTCTTCTCTCTCCACACCACTCTCT 149
Qy 36 SerSerProProPro-----ProSerThrThrThrThrThrThrThrThrThrThr 54
Db 150 TCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 209
Qy 55 ThrProSerAlaProProProSerProProThrProThrProThrProGlySerProPro 74
Db 210 CCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 257
Qy 75 LeuProGlnProSerProAlaProThrThrThrProGlySerProProAlaProValThr 94
Db 258 CTCTCTCAACCTTCCCGCTCCGCTCCCATCACTCTCT---TCTCCACCGTCTCCACCA 314
Qy 95 ProProThr---ArgAsnProProProSerValProGlyProProSerAsnProSerArg 113
Db 315 CCCTCAAAACCTCGAAGCCTCCATCTCTTAACCAAGGACCAACCAACACCTCCCTCA 371
Qy 114 GluGlyGlySerProArgProProSerSer-----ProSerProProSerProSerSer 131
Db 372 ---GGATCTACTCTAGAACCTCCATCAAACTAAACCGTCCGCGCGCTCTGATCTCTCC 428
Qy 132 AspGlyLeuSerThrGlyValValValGlyValAlaIleGlyValAlaLeuLeuVal 151
Db 429 GATGGATGTCACCGAGTTGTGTAGGATCCCATTCGCCATTGGAGAGTTCGCTATCTTGT 488
Qy 152 IleValThrLeuIleCysLeuLeuCysLysLysLysArgArgArg-----AspGluGlu 169
Db 489 ATACTGACTCTGATTTGTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 548
Qy 170 AspAlaTyrThrValProProProProProProProGlyProLysAlaGlyGlyProGly 189
Db 549 GCTGCTTACTATGTTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 608
Qy 190 GlyGlnGlnGlnGlnTrpArgGlnGlnAsnAlaThrProProSerAsp---HisValVal 208
Db 609 GGTCAACAACAGTATTGGCAACAACAACAACCGCTCCGCGCGCTCAGATAATCATGTAG 668
Qy 209 ThrSerLeuProProProProLysAlaProSerProProArgGlnProProProProPro 228
Db 669 ACGTCATTG---CCACCACCTTAAGCCTCCATCTCCACCAACCAACCAACCTCTCTCC 725

```

Qy 229 ProProPro---PheMetSerSerSerGlyGlySerAspTyrSerAspArgProValLeu 247
Db 726 CCACCACCAGCATTCATGAGTAGCAGTGGTCTGCTGACTATTCGGATCTTCGGTCTT 785
Qy 248 ProProProSerProGlyLeuValLeuGlyPheSerLysSerThrPheThrTyGluGlu 267
Db 786 CCTCCACCATTCTCAGGGCTTGTGTAGGCTTTCTTAAAGACACTTTTCACCTTATGAGGAG 845
Qy 268 LeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyPheGly 287
Db 846 TTGTCGAGAGCTACTAATGGCTTCTCTGAGGCTTAATTTGTAGGACAGGAGGGTTGGT 905
Qy 288 TyrValHisLysGlyValLeuProSerGlyLysGluValAlaValLysGlnLeuLysVal 307
Db 906 TATGTGCATATAAGTATATTCCTAGTGGGAAAGAGTTGCTGTGMAACAGTTCGAAGCT 965
Qy 308 GlySerGlyGlnGlyGluArgGluPheGlnAlaGluValGluLulleSerArgValHis 327
Db 966 GGTAGTGTCTCAGGGAGAGAGAGATTTCCAGGCTCAGGTTGAGATCAATAGCAGAGTTTCAT 1025
Qy 328 HisArgHisLeuValSerLeuValGlyTyrCysIleAlaGlyAlaLysArgLeuVal 347
Db 1026 CACAGGCAATTTGGTCTCTTATTTGGTTATTTGATGGCCGCTGTTCAAAGATTACTTGTCT 1085
Qy 348 TyrGluPheValProAsnAsnLeuGluLeuHisGlyGlyGlyArgProThr 367
Db 1086 TATGAGTTGTTCACCAACAATCTTGAGTTTCCCTCCATGTTAGGGAGCGCTACG 1145
Qy 368 MetGluTrpSerThrArgLeuLysIleAlaLeuGlySerAlaLysGlyLeuSerTyLeu 387
Db 1146 ATGGAATGGAGTACTAGATTGCTCTGCTGATCTGCTAAAGGACTTTCATATCTT 1205
Qy 388 HisGluAspCysAsnProLysIleLeHisArgAspIleLysAlaSerAsnIleLeu 407
Db 1206 CATGAAGATTCCAAATCCGAAATCATTCACCGTGATATTAGGGCTCAACATATTGATT 1265
Qy 408 AspPheLysPheGluAlaLysValAlaAspPheGlyLeuAlaLysIleAlaSerAspThr 427
Db 1266 GATTTCAAATTTGAAGCTAAGTTGCTGACTTTGCTCTGCTGCAAGATTGCTTCTGTATACA 1325
Qy 428 AsnThrHisValSerThrArgValMetGlyThrPheGlyTyrLeuAlaProGluTyAla 447
Db 1326 AACACTCATGTATCTACACGGCTGATGGAACTTTGGGTATTTGGCTCCGGAATATGCT 1385
Qy 448 AlaSerGlyLysLeuThrGluLysSerAspValPheSerPheGlyValValLeuGlu 467
Db 1386 GCAAGTGGAAAGCTCACAGAAAAGTCTGACGTTTTCTCATTTGGCGTTGTACTTTTGAA 1445
Qy 468 LeuIleThrGlyArgArgProValAspAlaAsnAsnValTyrValAspAspSerLeuVal 487
Db 1446 CTATTACTGGAGGCGCCCTGTTGATGCGAACATGTCTATGTATGATGACAGCTTAGTT 1505
Qy 488 AspTrpAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGluGlyLeuAla 507
Db 1506 GACTGGGCACGACCATTTGCTTAACCGAGCATCTCAGGAAGGAGATTTTGAGGGTTGGCT 1565
Qy 508 AspAlaLysMetAsnAsnGlyTyrAspArgGluGluMetAlaArgMetValAlaCysAla 527
Db 1566 GATTCAAGATGGGTAATAGTATGACAGAGAGAGATGGCTCCGATGGTTGCTTGGCT 1625
Qy 528 AlaAlaCysValArgHisSerAlaArgArgProArgMetSerGlnIleValArgAla 547
Db 1626 CGCGCTGTGTCCCATTCAGCTCGCGCAGACCTCGCATGACCGATAGTAGTACGGCG 1685
Qy 548 LeuGluGlyAsnValSerLeuSerAspLeuAsnGluGlyMetArgProGlyGlnSerAsn 567
Db 1686 TTAGAAGAAATGATCGCTGCTGATCTTAACGAAGGAGATGAGACCGGGTCAACGCAAC 1745
Qy 568 ValTyrSerSerTyGlyGlySerThrAspTyrAspSerSerGlnTyrAsnGluAspMet 587
Db 1746 GTATACAGCTCATATGAGGAAGCACAGACTATGACACGAGCCCAATCAACAGCAGCATG 1805

Qy 588 LysLysPheArgLysMetAlaLeuGlyThrGlnGluTyrAsnAlaThrGlyGlyTyrSer 607
Db 1806 ATAAAGTTTAGGAAAATGGCTCTTGGAACTCAAGATACGCACACCGGCGAGTACAGT 1865
Qy 608 AsnProThrSerAspTyrGlyLeuTyrProSerGlySerSerGlyGlnThrThr 627
Db 1866 AATCCAAACCACTGACTACGAGCTGTACCCGCTCTGGTTCAAGCAGTGAAGGTCACAGCCACA 1925
Qy 628 ArgGluMetGluMetGlyLysIleLeuArgThrGlyGlnGlyTyrSerGlyProSerLeu 647
Db 1926 CGAGAAATGAGATGGGAAAGATTAGAAAACCGTCAAGTTATAGTGGACCTCTCTT 1985
RESULT 11
BT008409 1959 bp mRNA linear PLN 15-MAY-2003
LOCUS Arabidopsis thaliana At3g24600 gene, complete cds.
DEFINITION Arabidopsis thaliana
ACCESSION BT008409
VERSION BT008409.1 GI:30725491
KEYWORDS FLI CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE 1 (bases 1 to 1959)
AUTHORS Cheuk,R., Chen,H., Kim,C.J., Shinn,P., Bowser,L., Carninci,P.,
Dale,J.M., Hayashizaki,Y., Heuan,V.W., Ishida,J., Jones,T.,
Kamiya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lin,J., Miranda,M.,
Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L.,
Sakurai,T., Satou,M., Seki,M., Southwick,A., Toriumi,M., Wong,C.,
Wu,H.C., Yamada,K., Yu,G., Yuan,S., Shinozaki,K., Davis,R.W.,
Theologis,A. and Ecker,J.R.
Arabidopsis ORF clones
Unpublished
2 (bases 1 to 1959)
AUTHORS Cheuk,R., Chen,H., Kim,C.J., Shinn,P., Bowser,L., Carninci,P.,
Dale,J.M., Hayashizaki,Y., Heuan,V.W., Ishida,J., Jones,T.,
Kamiya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lin,J., Miranda,M.,
Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L.,
Sakurai,T., Satou,M., Seki,M., Southwick,A., Toriumi,M., Wong,C.,
Wu,H.C., Yamada,K., Yu,G., Yuan,S., Shinozaki,K., Davis,R.W.,
Theologis,A. and Ecker,J.R.
Direct Submission
Submitted (15-MAY-2003) Salk Institute Genomic Analysis Laboratory
(SIGnAL), Plant Biology Laboratory, The Salk Institute for
Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
USA
RIKEN Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
Arabidopsis Full-length cDNA') : Seki,M., Narusaka,M., Ishida,J.,
Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
Hayashizaki,Y. and Shinozaki,K.
The Salk, Stanford, PGECC (SSP) Consortium members constructed and
sequenced the PUNI (ORF) clones using the RAFL cDNAs: Cheuk,R.,
Chen,H., Kim,C.J., Shinn,P., Bowser,L., Chan,M.M., Chang,C.M.,
Dale,J.M., Heuan,V.W., Jones,T., Karlin-Neumann,G., Lam,B.,
Lee,J.M., Lin,J., Miranda,M., Nguyen,M., Onodera,C.S., Palm,C.J.,
Quach,H.L., Southwick,A., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C.,
Yamada,K., Yu,G., Yuan,S., Davis,R.W., Theologis,A., and
Ecker,J.R.
Cheuk,R. (SSP/salk) and Seki,M. (RIKEN GSC) contributed equally to
this work. Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/salk)
contributed equally to this work as PIs.
FEATURES
source Location/Qualifiers
1. 1959
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="3"
/clone="U15366"
/ecotype="Columbia"


```

Qy 568 ValTyrSerSerTyrGlyGlySerThrAspTyrAspSerSerGlnTyrAsnGluAspMet 587
Db 1717 GTATACAGCTCATATGAGGAGACACAGCATATGACACGAGCAATACACGACGACATG 1776

Qy 588 LysLysPheArgLysMetAlaLeuGlyThrGlnGluTyrAsnAlaThrGlyGluTyrSer 607
Db 1777 ATAAAGTTTAAAGAAATGGCTCTTGGAACTCAAGAAATACGCGCACAAACGGCGAGTACAGT 1836

Qy 608 AsnProThrSerAspTyrGlyLeuTyrProSerGlySerSerGluGlyGlnThrThr 627
Db 1937 AATCAACCAAGTGAATGAGGACTGACCGCTCTGTTCAAGCAGTGAAGGTCAAGCCACA 1896

Qy 628 ArgGluMetGluMetGlyLysIleLysArgThrGlyGlnGlyTyrSerGlyProSerIleu 647
Db 1897 CGAGAAATGAGATGGAGAAAGATTAAAGAAACCGTCAAGGTTATAGTGGACCTCTCTT 1956

```

```

RESULT 12
AF370509 2257 bp mRNA linear PLN 30-APR-2001
LOCUS
DEFINITION
Arabidopsis thaliana protein kinase-like protein (MOB24.13) mRNA,
complete cds.
ACCESSION
AF370509.1 GI:13877616
VERSION
FLI CDNA.
SOURCE
Arabidopsis thaliana (thale cress)
ORGANISM

```

```

Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 2257)
Southwick,A., Karlin-Neumann,G., Nguyen,M., Lam,B., Miranda,M.,
Palm,C.J., Bowser,L., Jones,T., Banh,J., Carninci,P., Chen,H.,
Cheuk,R., Chung,M.K., Hayashizaki,Y., Ishida,J., Kamiya,A.,
Kawai,J., Kim,C., Lin,J., Liu,S.X., Narusaka,M., Pham,P.K.,
Sakano,H., Sakurai,T., Satou,M., Seki,M., Shinn,P., Yamada,K.,
Shinozaki,K., Ecker,J., Theologis,A. and Davis,R.W.

```

REFERENCE

```

AUTHORS
Direct Submission
Submitted (17-APR-2001) DNA Sequencing and Technology Center,
Stanford University, 955 California Avenue, Palo Alto, CA 94304,
USA
COMMENT
e-mail for correspondence: arab@sequence.stanford.edu

```

RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA: 'RIKEN Arabidopsis Full-length cDNA'): Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGEC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Southwick,A., Karlin-Neumann,G., Nguyen,M., Lam,B., Miranda,M., Palm,C.J., Bowser,L., Jones,T., Banh,J., Chen,H., Cheuk,R., Chung,M.K., Kim,C., Lin,J., Liu,S.X., Pham,P.K., Sakano,H., Shinn,P., Yamada,K., Ecker,J., Theologis,A. and Davis,R.W.

Southwick,A. (SSP/Stanford) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Davis,R.W. (SSP/Stanford) contributed equally to this work as PIs.

FEATURES

source

```

Location/Qualifiers
1. .2257
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="III"
/clone="RAFL11-07-N23"
/ecotype="Columbia"
/notes="This clone is in pBluescript"
1. .2257
/gene="MOB24.13"
107. .2065
/gene="MOB24.13"
/codon_start=1

```

gene

CDS

```

/product="protein kinase-like protein"
/protein_id="AAK43886.1"
/db_xref="GI:13877617"
/translation="MSTAPSPGTTTPSPSPPTNSPTTTPPAASPPPTPTTSSPP
PSSTNSTSPSPSPPLPSPSPSPGSLTPPIQPSPSPAPITPSPSPPTSPSPSP
PSNQPPNTSPSGTPTPTSPNTKPSPPSSDGLSTGVVGIAGVAILVILTLICL
LCKKRRRRHDDAAAYVPPPPSPGKAGPYGQQQYQQQNSPSPDHVVTSLPP
PKPSPRPKPPPPPPAFMSSSGSDYSDLPVLPSPGLVLFSGSTTTEYLSRA
TNGFSBANLGGQGGFVGHKGLPSPGKEVAVKOLKAGSGGGERFOAEVETISRHH
HLVSLICYMAGVQRLVLFVFPNNLEFHLHGKGRPTMESTRLKALGSAKGLSYL
HEDCNPKIHRDIKASNILLIDPFEAKVADPGLAKTASDNTNTHSVTRMGTFGYLAPE
YASGKLTKEKDFVSFGVLELITGRPDANNVYVDDSLDWARLPULLRARSSEGF
EGLASKMGNEYDREEMARVACAAACVRSARRRPMQSVLRLEGNVLSLDSLEGM
RPGHSNVSYSGGSTDYDTSQYNDMDIKFRKMAIGTQEYGTGTEYSNPTSDYGLYPSG
SSSEGCATREMENGKIKTKQYSGPSL"

```

ORIGIN

```

Alignment Scores:
Pred. No.: 1.52e-63 Length: 2257
Score: 2934.50 Matches: 573
Percent Similarity: 90.61% Conservative: 25
Best Local Similarity: 86.82% Mismatches: 41
Query Match: 84.98% Indels: 21
DB: 8 Gaps: 12

US-10-086-464-2 (1-647) x AF370509 (1-2257)

Qy 1 MetSerSerAlaProSerProGlyThr-----GlySerProProSerProSer 17
Db 107 ATGTCCACACGCGCGTCTCCAGGACATCTTCATCATCATCTCCACGCTCTCTCC 166

Qy 18 AsnSerThrThrThrProProAla-----SerAlaProProThrThrPro 35
Db 167 AATCGACACACCACTCTCTCTCCAGAGCGCTCTCTCTCTCTCCACCACTCTCT 226

Qy 36 SerSerProProPro-----ProSerThrIleProThrSerProProSerSer 54
Db 227 TCTTCTCTCTCTCGTCCGCGTCAACTAATCAACCTCTCTCTCTCTCTCTCTCTTA 286

Qy 55 ThrProSerAlaProProProProProProProProProProProProProPro 74
Db 287 CTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 334

Qy 75 LeuProGlnProSerProProAlaProThrThrProGlySerProProAlaProValThr 94
Db 335 ATTCTCTCAACTTCTCCCTCGCTCCATCATCTCTCTCTCTCTCTCTCTCTCTCT 391

Qy 95 ProProThr---ArgAsnProProProSerValProGlyProProSerAsnProSer 113
Db 392 CCTCAAAACCTCGAAGCCTCCATCTCTAACAAGGACCAACCAACACTCTCTCTCA 448

Qy 114 GluGlyGlySerProProProProProSer-----ProSerProProSerProSer 131
Db 449 ---GGATTTACTCTCTGAATCTCAACACTAAACCGTCCGCGCGCTCTGATCTTCC 505

Qy 132 AspGlyLeuSerThrGlyValValValGlyIleAlaIleGlyGlyValAlaLeuVal 151
Db 506 GATGGATTGTCTACCGGAGTTGTTGGTAGAATGCCATTGGAGGAGTCGCTATTCTGTT 565

Qy 152 IleValThrIleLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 169
Db 566 ATACTGACTCTGATTTGTCTCTCTGTAGAAGAAAGAAAGAACACACGACGATGAA 625

Qy 170 AspAlaTyrTyrValProProProProProProProProProProProProPro 189
Db 626 GCTGCTTACTATGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 685

Qy 190 GlyGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 208
Db 686 GTTCAACACAGTATTGGGCAACCAACCAACCGTTCACGCGCGCTCATGATCATG 745

Qy 209 ThrSerLeuProProProProLysAlaProSerProProArgGlnProProPro 228

```


Db 746 ACGTCATTG---CCACCACCTAAGCCTCCATCTCCACCAACGAAACCTCTCCGCCACCT 802

Qy 229 ProProProSerSerSerGlyGlySerAspTyrSerAspArgProValLeu 247

Db 803 CCACCACGACCATTCATGAGTAGCAGTGGTGGTCTGACATTATCCGATCTTCCGGTCTT 862

Qy 248 ProProProSerProGlyLeuValLeuGlyPheSerLysSerThrPheThrTyGluGlu 267

Db 863 CCTCCACCATCTCCAGGCGCTGTGTAGGCTTTCTTAAAGCACTTTCACTTATGAGGAG 922

Qy 268 LeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyPheGly 287

Db 923 TTGTGAGAGACTACTAATGGCTTCTCTGAGGCTTAATTTGTAGCAACAGGAGGTTGGT 982

Qy 288 TyrValHisLysGlyValLeuProSerGlyLysGluValAlaValLysGlnLeuLysVal 307

Db 983 TATGTGCATAAAGCTATATAGCTAGTGGGAAAGAGTTGCTGTGAAACAGTTCGAAAGCT 1042

Qy 308 GlySerGlyGlnGlyGluArgGluPheGlnAlaGluValGluIleLysArgValHis 327

Db 1043 GGTAGTGGTCAGGAGAGAGAGAGTTCAGGCTGAGGTTGAGATCATTTAGCAGAGTTTCAT 1102

Qy 328 HisArgHisLeuValSerLeuValGlyTyrCysIleAlaGlyAlaLysArgLeuVal 347

Db 1103 CACAGGCAATTTGGTTTCTCTTATTTGGTATTGTATGGCGGTGTCAAGATTTACTTGTCT 1162

Qy 348 TyrGluPheValProAsnAsnLeuLeuGluLeuHisLeuHisGlyGlyArgProThr 367

Db 1163 TATGAGTTGTTTCCAAACAACAATCTTGAGTTTCCCTCCATGTTAAGGACGCGCTACG 1222

Qy 368 MetGluTrpSerThrArgLeuLysIleAlaLeuGlySerAlaLysGlyLeuSerTyLeu 387

Db 1223 ATGGAATGGAGTACTAGATTGAAGATTGCTCTTGATCTGCTAAAGGACTTTTCATATCTT 1282

Qy 388 HisGluAspCysAsnProLysIleIleHisArgAspIleLysAlaSerAsnIleLeuIle 407

Db 1283 CATGAGATTGCATTCGAAATCATTCACCGTATATTAAAGGCTCAACATTTAGATT 1342

Qy 408 AspPheLysPheGluAlaLysValAlaAspPheGlyLeuAlaLysIleAlaSerAspThr 427

Db 1343 GATTTCAAATTTGAAGCTAAGGTTGCTGACTTTGCTTGTCCCAAGATTGCTTCTGTATACA 1402

Qy 428 AsnThrHisValSerThrArgValMetGlyThrPheGlyTyrLeuAlaProGluTyAla 447

Db 1403 AACACTCATGTATCTACACGGCTGATGGAAACCTTTGGGTATTTGGCTCCGGAATATGCT 1462

Qy 448 AlaSerGlyLysLeuThrGluLysSerAspValPheSerPheGlyValValLeuLeuGlu 467

Db 1463 GCAAGTGGAAAGCTCACAGAAAGCTGACGTTTCTCATTTGGCGTTGTACTTTGGAA 1522

Qy 468 LeuIleThrGlyArgArgProValAspAlaAsnValTyrValAspAspSerLeuVal 487

Db 1523 CTTATTACTGGAGCGCGCCTGTTGATGCGAAACATGCTATGTATGATGACAGCTTAGTT 1582

Qy 488 AspTrpAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGluGlyLeuAla 507

Db 1583 GACTGGGCACACCATTTGCTTAAACCGACATCTCAGGAAGAGATTTTGAGGGGTTTGCT 1642

Qy 508 AspAlaLysMetAsnAsnGlyTyrAspArgGluGluMetAlaArgMetValAlaCysAla 527

Db 1643 GATTCAAAGATGGGTATGATGATGACAGAGAGAGATGGCTCCGATGGTTGCTTGCGCT 1702

Qy 528 AlalaCysValArgHisSerAlaArgArgProArgMetSerGlnIleValArgAla 547

Db 1703 GCGGCTTGTTGCCATTTACGTCGCGCAGACCTCGCATGAGCCAGATAGTACGGGCG 1762

Qy 548 LeuGluGlyAsnValSerLeuSerAspLeuAsnGluGlyMetArgProGlyGlnSerAsn 567

Db 1763 TTGAAGGAAATGATCGCTCTGATCTTAACGAAGGATGAGACCGGGTCACAGCAAC 1822

Qy 568 ValTyrSerSerTyrGlyGlySerThrAspTyrAspSerSerGlnTyrAsnGluAspMet 587

Db 1823 GTATACAGCTCATATGGAGGAGACACAGACTATGACACGACCAATACACAGCACATG 1882

Qy 588 LysLysPheArgLysMetAlaLeuGlyThrGlnGluTyrAsnAlaThrGlyGlyTyrSer 607

Db 1883 ATAAAGTTTATAGAAAATGGCTCTTGGAACTCAAGAATACCGCACACCGCGAGTACAGT 1942

Qy 608 AsnProThrSerAspTyrGlyLeuTyrProSerGlySerSerSerGlyGlyGlnThrThr 627

Db 1943 AATCCAACCAAGTACTACGAGCTGTACCGCTCTGGTTCAAGCAGTGAAGTCAAGCCACA 2002

Qy 628 ArgGluMetGluMetGlyLysIleLysArgThrGlyGlnGlyTyrSerGlyProSerLeu 647

Db 2003 CGAGAAATGAGATGGGAAAGATTAGAAAACCGTCAAGGTTATAGTGGACCTCTCTT 2062

RESULT 13

AY089024 2324 bp mRNA linear PLN 14-APR-2003
LOCUS Arabidopsis thaliana clone 17909 mRNA, complete sequence.

DEFINITION AY089024

ACCESSION AY089024

VERSION AY089024.1 GI:21407798

KEYWORDS Full cDNA.

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM

Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

REFERENCE

AUTHORS

TITLE

Genome Biol. 3 (6), RESEARCH0029 (2002)

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

This clone sequence is one of 5,000 Ceres full-length cDNAs made of available to TIGR and Genbank. The following quality assessment of this set was done by comparison with known proteins: two percent of the clones are estimated to be 5'-truncated; less than one percent are 3'-truncated; approximately two percent represent alternative splice variants, including unspliced introns and spliced exons; one percent may contain premature stop codons; five percent may have frame shifts in a coding region. A sequence is considered to be 5'-truncated if it lacks the translation initiation start (ATG). A sequence is considered to be 3'-truncated if it lacks the C-terminal end of the encoded protein. Please note that these cDNA sequences are derived from the Ws or Laer ecotypes and therefore may contain polymorphisms when compared to sequences from Col-0. Genset carried out the library production and sequencing of the full-length clones. Ceres, Inc. carried out the clustering of the 5' sequences, selection of clones, and sequence assembly.

FEATURES

source

1. .2324
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="17909"

ORIGIN

Alignment Scores: 2.45e-62 Length: 2324
Pred. No.: 2885.00 Matches: 570
Percent Similarity: 90.30% Conservative: 26
Best Local Similarity: 86.36% Mismatches: 42

Query Match: 83.55% Indels: 23
DB: 8 Gaps: 13
US-10-086-464-2 (1-647) x AY089024 (1-2324)

Qy 1 MetSerSerAlaProSerProGlyThr-----GlySerProProSerProProSer 17
Db 105 ATGTCACAGCGCGCTCCAGGCACTACTCCATCACCCTCTCCACCGCTCTCTCCACCA 164
Qy 18 AenSerThrThrThrProProAla-----SerAlaProProProThrPro 35
Db 165 AACTCGACAACCACTCTCTCCAGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 224
Qy 36 SerSerProProPro---ProSerThrThrThrProProProProSerSerArgSer 54
Db 225 TCTTCT 284
Qy 55 ThrProSerAlaProProProSerProProThrThrProProSerProGlySerProPro 74
Db 285 CCT 332
Qy 75 LeuProGlnProSerProAlaProThrThrProGlySerProProAlaProValThr 94
Db 333 CTCTCTCAACCTTCTCAATCGCTCCCATCACTCTCT---TCCCGCGCTCTCTCCACCA 389
Qy 95 ProProThr---ArgAsnProProProSerValProGlyProProSerAsnProSerArg 113
Db 390 CCTCAAAACCTCGAAGCCCTCTCTCTTAACCAAGGACCAACCAACCTCTCTCTCA 446
Qy 114 GluGlySerProArProProSerSer-----ProSerProProProSerProSer 131
Db 447 ---GGACTACTCTAGAACTCCATCAAAACGCTAAACGCTCGCCACCGCTCTGATTTCC 503
Qy 132 AspGlyLeuSerThrGlyValValValGlyLeuAlaIleGlyValAlaLeuVal 151
Db 504 GATGATTGTCTACCGAGTGTGTGTAGAAATCGCCATTTGGAGGAGTCTGATTTCTGTT 563
Qy 152 IleValThrLeuIleCysLeuLeuCysIleValValValValValValValValVal 169
Db 564 ATACTGACTCTGATTTGTCTCTCTGTAAAGAAACCAAGACAGACAGATGAA 623
Qy 170 AspAlaTyThrValProProProProProProGlyProProGlyProLysAlaGlyProTyGly 189
Db 624 GCTGTTACTATGTT---CCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 680
Qy 190 GlyGlnGlnGlnInTrpArgGlnGlnAsnAlaThrProProSerAsp---HisValVal 208
Db 681 GGTCAACAACAGTATTGGCAACAAACAAACCGCTCACGGCGCTCAGATAATCATGTAGT 740
Qy 209 ThrSerLeuProProProProLysAlaProSerProProArgGlnProProProProPro 228
Db 741 ACCTCATG---CCACCACTTAAGCTCTCATCTCCACCAACCAACCTCTCTCCGCCACT 797
Qy 229 ProProPro---PheMetSerSerSerGlyGlySerAspTyThrSerAspArgProValLeu 247
Db 798 CCACCACAGCATTCATCAGTAGCAGTGGTCTGTTCTGACTATTCTGACTCTCTCTCTCT 857
Qy 248 ProProProSerProGlyLeuValLeuGlyPheSerIleValSerThrThrThrThrGlu 267
Db 858 CCTCCACCATCTCCAGGGCTTGTGTAGGCTTTTCAAAAACCACTTTTCACTTATAGAGAG 917
Qy 268 LeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyPheGly 287
Db 918 TTGTCAAGAGTACTAATGGCTTCTCTGAGGCTAATTTGTAGCAACAGGAGGTTTGGT 977
Qy 288 TyrValHisIleGlyValLeuProSerGlyLysGluValAlaValLysGlnLeuLysVal 307
Db 978 TATGTGCATAAAGGTATATTCCTAGTGGGAAAGAGTGTGTGTAACAGCTTGAAGCT 1037
Qy 308 GlySerGlyGlnGlyGluArgGluPheGlnAlaGluValGluIleIleSerArgValHis 327
Db 1038 GGGAGTGTCTAGGAGAGAGAGGTTTCAGGCTGAGGTTGAGTCAATTAGCAGAGTTTCAT 1097

Qy 328 HisArgHisLeuValSerLeuValGlyTyThrCysIleAlaGlyAlaLysArgLeuVal 347
Db 1098 CACAGGCAATTTGGTTTCTTATTGGTTATTGTATGCTGATGTTCAAGATGCTTGTCT 1157
Qy 348 TyrGluPheValProAsnAsnAsnLeuGluLeuHisIleHisGlyGluArgProThr 367
Db 1158 TATGAGTTTGTTCCTAAACAATCTTGAAGTTTCACTCTCCATGTTAAGGAGGCTTACA 1217
Qy 368 MetGluTrpSerThrArgLeuLysIleAlaLeuGlySerAlaLysGlyLeuSerTyLeu 387
Db 1218 ATGAAATGGAGTACTAGATTGAAGATTCTCTTGATCTCTGTAAGGACTTTCATATCT 1277
Qy 388 HisGluAspCysAsnProLysIleIleHisArgAspIleLysAlaSerAsnIleLeu 407
Db 1278 CATGAAGATTGCAATCCGAAATCATTCACCGTGATTAAGGCGTCAACATATTGATT 1337
Qy 408 AspPheLysPheGluAlaLysValAlaAspPheGlyLeuAlaLysIleAlaSerAspThr 427
Db 1338 GATTTCAAATTTGAAGCTAAGGTTGCTGACTTTGGTCTTGCAGGATTTGCTTCTGATA 1397
Qy 428 AsnThrHisValSerThrArgValMetGlyThrPheGlyTyThrLeuAlaProGluTyAla 447
Db 1398 AACACTCATGTATCTACCGCTGATGGAACTTTGGGTATTGGCTCCAGAATATGCT 1457
Qy 448 AlaSerGlyLysLeuThrGluLysSerAspValPheSerPheGlyValValLeuGlu 467
Db 1458 GCAAGCGAAAGCTCACAGAAAGCTCTGACGTTTCTCAITTTGTTGTTACTTTGGAG 1517
Qy 468 LeuIleThrGlyArgArgProValAspAlaAsnAsnValTyValAspAspSerLeuVal 487
Db 1518 CTTATAACCGGAGGCGCCCTGTTGATGCAACAATCTCTATGTAGATGACAGCTTAGT 1577
Qy 488 AspTrpAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGlyLeuAla 507
Db 1578 GACTGGGACACGACCTTGTACCGAGCATCTGAGGAGGAGATTTTGTATGGTTGGCT 1637
Qy 508 AspAlaLysMetAsnAsnGlyTyThrAspArgGluMetAlaArgMetValAlaCysAla 527
Db 1638 GATTCAAGATGGGTATAGTATAGTAGAGAGGAGATGGCTCGCATGTTGCTTGGCT 1696
Qy 528 AlaAlaCysValArgHisSerAlaArgArgProArgProMetSerGlnIleValArgAla 547
Db 1697 GCAGCTTGTGTTCGCCATTCAGCTCGCGCAGACCTCGCATGAGCCAGATAGTACGGCG 1756
Qy 548 LeuGluGlyAsnValSerLeuSerAspLeuAsnGluGlyMetArgProGlyGlnSerAsn 567
Db 1757 TTAGAAGGAATGTATCGTCTGTCTTAAACGAGGAGATGAGACCGGGTCAAGCAAC 1816
Qy 568 ValTySerSerTyThrGlySerThrAspTyThrAspSerSerGlnTyAsnGluAspMet 587
Db 1817 GTATACAGCTCATATGGAGGAGGAGCATGACTATGACACGAGGCAATACAGATGACATG 1876
Qy 588 LysLysPheArgLysMetAlaLeuGlyThrGlnGluTyThrAsnAlaThrGlyGlyTySer 607
Db 1877 AAGAAGTTTAGGAAATGGCTCTTGGAACTCAAGATACGCGCACACCGCGGAGTACAGT 1936
Qy 608 AsnProThrSerAspTyThrGlyLeuTyThrProSerGlySerSerSerGluGlyThrThr 627
Db 1937 AATCCAAACAGTACTACGAGCTGTACCCGCTCTGGTTCAAGCAGTGAAGGTCAAGCCACA 1996
Qy 628 ArgGluMetGluMetGlyLysIleLysArgThrGlyGlnGlyTyThrSerGlyProSerLeu 647
Db 1997 CGAGAAATGAGATGGGAAAGATTAAAGAAACCGTCAAGGTTATAGTGGACCTCTCT 2056

RESULT 14

AB020746
LOCUS
DEFINITION Arabidopsis thaliana genomic DNA, chromosome 3, P1 clone: MOB24.
ACCESSION AB020746 BA000014
VERSION AB020746.1 GI:3985949
KEYWORDS
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE

AUTHORS

TITLE

Kaneko,T., Katoh,T., Sato,S., Nakamura,A., Asamizu,E. and Tabata,S.
Structural analysis of Arabidopsis thaliana chromosome 3. II.
Sequence features of the 4,251,695 bp regions covered by 90 P1, TAC
and BAC clones
DNA Res. 7 (3), 217-221 (2000)

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

2 (bases 1 to 79706)
Kaneko,T., Katoh,T., Sato,S., Nakamura,Y., Asamizu,E. and Tabata,S.
Direct Submission
Submitted (03-DEC-1998) Yasukazu Nakamura, Kazusa DNA Research
Institute, Department of Plant Gene Research; 1532-3, Yana,
Kisarazu, Chiba 292-0812, Japan (E-mail: ynakamu@kazusa.or.jp,
Tel:81-438-52-3935, Fax:81-438-52-3934)

COMMENT

Address for correspondence: kaos@kazusa.or.jp
For the latest information on annotation of this clone, please see
http://www.kazusa.or.jp/kaos/cgi-bin/aggd graph.cgi?c=MOB24
Genes with similarity to proteins in the databases are described in
'product' or 'note' qualifiers. Genes that have no significant
protein similarity are described as 'unknown protein'.
The software programs used to predict genes include: Grail
(Informatics Group, Oak Ridge National Laboratory,
http://compbio.ornl.gov/Grail-1.3/),
GENSCAN (Chris Burge, MIT, http://CCR-081.mit.edu/GENSCAN.html),
NetGene2 (S.M. Hebsgaard, et al., CBS, Technical University of
Denmark, http://www.cbs.dtu.dk/services/NetGene2/) and
SplicePredictor (Volker Brendel, Stanford University,
http://genome.zoology.ualberta.edu/cgi-bin/sp.cgi).
Genes encoding tRNAs are predicted by tRNAscan-SE
(Sean Eddy, Washington University School of Medicine, St. Louis,
http://genome.wustl.edu/eddy/tRNAscan-SE/).
This sequence may not be the entire insert of this clone. It may be
shorter because we remove overlaps between neighboring submissions.
The 5' clone is MKP5 and the 3' clone is MSD24.

FEATURES

source

1..79706
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia"
/db_xref="taxon:3702"
/chromosome="3"
/clone="MOB24"
/clone_lib="Mitsui P1"
2398..3399
/note="gene_id:MOB24.1
unknown protein"

CDS

/codon_start=1
/evidence=not experimental
/protein_id="BAB01995.1"
/db_xref="GI:9294038"
/translation="MGDSSEDTGYPKFYSLNQHPMYSRPDKHAYNEEDED
VEGEDPQGGTIRNERPQKRNKQVSGFEPAGSDAKVADNRQEPALVLEW
GDRFQLGRSLRNDNEVAEKSEELRMEKSETQCRMRDIDLKRYKRYKKEVKS
GLSSKWSFFNNKLLMLCCVSPKSDGLACGVDSGEFVPMNTKVYLDKSGPDWMDSP
GDSSEDDDEVEYERKRVNDAASYKMLADSVRFKGYEKMEKSKKEQMELEKMR
ADFQDLQKQIIVDRAQSEIARLREBEENHGGDDDESEDEMNDSVNLSD"
complement(join(3794..4276,4363..4490,4577..4673,
4757..5179,5354..5548,5693..5970,6088..6214,6297..6496,
6601..6648,6743..6776,6882..7013,7464..7559,7677..7743,
8162..8217,8290..8385,8470..8998,9122..9462))

CDS

/codon_start=1
/evidence=not experimental
/product="DNA repair protein Muts"
/protein_id="BAB01996.1"
/db_xref="GI:9294039"
/translation="MQRQSRILSFQKPTATTGVLGSDAASGGGGGPRFNVNKEG
DAKGDSVRFAVSKSVDRGTDTPEKVPVRVLPSPGFKPAESAGDASSLSFNIMHKF
VKVDRDCSGERSREDVPLNDSSLCMKANDVIQPRSNNGKTQRNHAFSFGRAEL

RSVEDIGVDGVQGPETPGMRPRASRLKRVLEDMTFKEDKVPVLDSNKRLLKMLQDPV
CGKEKVEGTKEFWLESSRIDANRRRPPDPLYDKRTLHIPPDPVFKMSASQKQYMS
KLEAYMDIVLFFKVKGEFELYELDAELGHELDWMKMTSGVKCRQKVGISSEGDYAV
QKLIARGYKVRIEQLETSQAKARGANTTIPKLVQLVLPSTASEGNIQPDVHLLA
IKETKMLQKSCVTYGFADCAALRFWVGSISDDASCAALGALLMQVSKSEVLYDSK
GLSREACQALRYTLTGSTAVOLAPVQVMGDTDAAGVRNIIESNGYFKGSSSESWNCA
VDGNECDVALSALGELINHLRLKLELDVLKHGDIIPYQVVRGCLRIDGOTMNVLEIF
NNCDGPGSTLYKYLNCVSPGKRLLRNWI CHPLKDVESINKRLDVREFTANSES
MQITQYHLKLPDLERLGRISKSSVRSSASVLPALLGKVKLKQKVPAGQVLKGRFSG
IDLALLQKESNMMSLLYKCLPILVKGSGLELFSQFAADIDFPNTQVQNDTE
NAETLTILIEFIERATQWSEVIHTISCLDVLRSAFAAASLSAGSMARVPPFSEAT
GOKTKGPILKIQGLWHPAVALGOLPVPNDILLGEARRSSGSIHPRSLITGPNM
DGKSTLRATCLAVIFAQGCYVPCSECEISLVDTITFRLGASDRIMTGESTFTVEPT
ETASVLNATQDSLVIIDELGRGTSFDGVAIAYSVRHLVEKVCQCMLEFATHVPIIT
KFNASHPRVTSKHWACAFKRSQDPQPGCCQDLVFLRLTEGACPESGLQVALMAGI
PNQVETASGAQAQMKRSIGENPKFSSELSRSLHEDWLKSLVGISRVAHNNAPIGE
DDVDTLFLWHEIKSSYCVPK"

CDS

10131..10577
/note="gene_id:MOB24.3"
/codon_start=1
/evidence=not experimental
/product="ethylene-responsive transcriptional
coactivator-like protein"
/protein_id="BAB01997.1"
/db_xref="GI:9294040"
/translation="MPSRYPGANTQDMPEPVVLHKSQKQSDLRDPKAVNALRNGVAV
QTVKFDAGNKGKSTAVPINTKLUKEEETEPAMDVRKAEVRLMTQKARLEKMSQ
ADAKQINERTQVQVEYENAGVNPQAVLAKMEKVLGVLRGKIGK"
complement(join(11108..11381,11467..11528,11699..11836,
11934..12071,12161..12334,12421..12510,12718..13104,
13623..13765,14306..14405))
/note="gene_id:MOB24.4"
/codon_start=1
/evidence=not experimental
/product="aldehyde dehydrogenase"
/protein_id="BAB01998.1"
/db_xref="GI:9294041"
/translation="MENGKNGATTVKLPEIKFTKLFINGOFIDAAASKTETIDPRN
GEVIATAEKGDVDAVNAARYAFAGHPWRMTGFERAKLKNKFDLRIENIELBA
KLDAVDGKGLFQLGKYADIATAGHFRYNAGADKHGHTKMTFRSLFGYTLKEPIG
VVGNIIPWNPFSIMFATKAPAMAAGCTVWVPAEQTSLSALFYAHLSEKAGIPDGLV
NIVTFGSTAGAAIAASHMDVKVSTGTSVGRKIMOAAAASNLKVSLELGGKSPIL
IFNDADIDKAADLALGCFYNKGEICVASSRVFQEGIDKVKVEKLEKADMTVGD
FDSTARQGPQVKKQFKILSYIEHGNEGATLLTGKKAIDGKGYFIQPTIFADVTED
MKIYDEIFGVSMLMKFTVEBIGKANNKYGLAAGILSQDIDLINTVSRSLKAGI
IWNVCYFGLDCCPYGGYKMSGNCRESGMDALDNYLQTSVVMPLHNSPMW"

CDS

complement(join(19742..19933,20400..20657))
/note="gb|AAD25142.1
gene_id:MOB24.5
similar to unknown protein"
/codon_start=1
/evidence=not experimental
/protein_id="BAB01999.1"
/db_xref="GI:9294042"
/translation="WTPLSIASPSPWPLLSPPLSSSSNLRAAPSIFFSPSRINHGF
SCSFIQVRNALSRTKQLQVAVMAPEEKLTRNPLDPEIWERPKGRRPDIFFPK
FSPMKSFLPPMPYDPPEEEEBEKEETEDDPEKEDQPNRS"
join(23402..23459,23819..24021)
/note="gene_id:MOB24.6
unknown protein"
/codon_start=1
/evidence=not experimental
/protein_id="BAB02000.1"
/db_xref="GI:9294043"
/translation="MQNASLKLPLLIFILVITSNLGAEARKLTGVADVIVEGEAASPO
YGIIDENDTLNKSCKRDICTQCCRGGFCNLILQKCECL"
complement(join(25945..26640,26714..26824,26914..27012,
27119..27224,27291..27360,27436..27583))
/note="contains similarity to ubiquitin conjugating enzyme
gene_id:MOB24.7"
/codon_start=1
/evidence=not experimental
/protein_id="BAB02001.1"

CDS

similar to unknown protein"
/codon_start=1
/evidence=not experimental
/protein_id="BAB01999.1"
/db_xref="GI:9294042"
/translation="WTPLSIASPSPWPLLSPPLSSSSNLRAAPSIFFSPSRINHGF
SCSFIQVRNALSRTKQLQVAVMAPEEKLTRNPLDPEIWERPKGRRPDIFFPK
FSPMKSFLPPMPYDPPEEEEBEKEETEDDPEKEDQPNRS"
join(23402..23459,23819..24021)
/note="gene_id:MOB24.6
unknown protein"
/codon_start=1
/evidence=not experimental
/protein_id="BAB02000.1"
/db_xref="GI:9294043"
/translation="MQNASLKLPLLIFILVITSNLGAEARKLTGVADVIVEGEAASPO
YGIIDENDTLNKSCKRDICTQCCRGGFCNLILQKCECL"
complement(join(25945..26640,26714..26824,26914..27012,
27119..27224,27291..27360,27436..27583))
/note="contains similarity to ubiquitin conjugating enzyme
gene_id:MOB24.7"
/codon_start=1
/evidence=not experimental
/protein_id="BAB02001.1"

CDS

similar to unknown protein"
/codon_start=1
/evidence=not experimental
/protein_id="BAB01999.1"
/db_xref="GI:9294042"
/translation="WTPLSIASPSPWPLLSPPLSSSSNLRAAPSIFFSPSRINHGF
SCSFIQVRNALSRTKQLQVAVMAPEEKLTRNPLDPEIWERPKGRRPDIFFPK
FSPMKSFLPPMPYDPPEEEEBEKEETEDDPEKEDQPNRS"
join(23402..23459,23819..24021)
/note="gene_id:MOB24.6
unknown protein"
/codon_start=1
/evidence=not experimental
/protein_id="BAB02000.1"
/db_xref="GI:9294043"
/translation="MQNASLKLPLLIFILVITSNLGAEARKLTGVADVIVEGEAASPO
YGIIDENDTLNKSCKRDICTQCCRGGFCNLILQKCECL"
complement(join(25945..26640,26714..26824,26914..27012,
27119..27224,27291..27360,27436..27583))
/note="contains similarity to ubiquitin conjugating enzyme
gene_id:MOB24.7"
/codon_start=1
/evidence=not experimental
/protein_id="BAB02001.1"

CDS

similar to unknown protein"
/codon_start=1
/evidence=not experimental
/protein_id="BAB01999.1"
/db_xref="GI:9294042"
/translation="WTPLSIASPSPWPLLSPPLSSSSNLRAAPSIFFSPSRINHGF
SCSFIQVRNALSRTKQLQVAVMAPEEKLTRNPLDPEIWERPKGRRPDIFFPK
FSPMKSFLPPMPYDPPEEEEBEKEETEDDPEKEDQPNRS"
join(23402..23459,23819..24021)
/note="gene_id:MOB24.6
unknown protein"
/codon_start=1
/evidence=not experimental
/protein_id="BAB02000.1"
/db_xref="GI:9294043"
/translation="MQNASLKLPLLIFILVITSNLGAEARKLTGVADVIVEGEAASPO
YGIIDENDTLNKSCKRDICTQCCRGGFCNLILQKCECL"
complement(join(25945..26640,26714..26824,26914..27012,
27119..27224,27291..27360,27436..27583))
/note="contains similarity to ubiquitin conjugating enzyme
gene_id:MOB24.7"
/codon_start=1
/evidence=not experimental
/protein_id="BAB02001.1"

```
/db_xref="GI:9294044"
/translation="MAQARLSRMOKELKLLGDPHPHAGSPHLSAANGSGDFSTF
STIDAOLEGPDVYANGIENLKIQIPERYFPQPIIVSFATPIVHPNIDNSGRICLDI
LNLPPKQWQSLNISTVLTSRLLSEPDDGLMCEVSREYKXNQTFDYKAREMT
EDNGKRAVVGFEGNSFGNDIKRKKLSLALPSQKQKJCGEELTRGVSAAC
KENKPNLNGKLSLGLKQPCNDTLASPTSAKSDNNRSLRKLRLSPLGELNEVSK
PEVLAQDTMKLEMNQEDARSURGFENSFVLETSMAESIIVLSDSDSGQEBEERVS
SRSLSLAKRRVLAKRP"
```

CDS

```
28636..30168
/note="emb|CAB79678.1
gene id:MOB24.8
similar to unknown protein"
```

```
/codon_start=1
/evidence=not experimental
```

```
/protein_id="BAB02002.1"
```

```
/db_xref="GI:9294045"
```

```
/translation="MSSLIAYLSFNSFPLKLSRILSNLWLVRTSTNKKRWCL
AIPTYHLHATPLWKSSITMLLRQSQMVPEPTTYFEISLGRAFILQAMGHGETLLWD
TTSFLGTRILWFSSASPHQONQSTTFSPENDASLSPPYNSRHHDLHLSEGLK
GLQVLGLVFGPLRYRPIISKDDPKLPLKPRDLNYSLSIFQSLSVHLLASIDRHH
TNHNLVLIWLKNSLRTPLVSSRTFVTCGICLKNGIMPEPTTSRWCDLCHPSATI
CISTHLGISKQSCISRVKASISKEVPNLFPIITPSPHALVVGKIKIGIMI
TSPSGDYQRFNLSHYSPLITEPKQKRFNLIKFSIHDLQSLKNGIMISPMWDDVR
SFFNLDLPTIVTHLGDQQLLVARSARTSALEPYSTSRLLITVTNLLSIDSFVED
PSTTRDLTCAIMPYSLSLWKLALNFMWSFHIIVHLVLGNASPNYVNLNFIPESLILIFV
LI"
```

CDS

```
join(32831..33058,33143..33907)
```

```
/note="contains similarity to heat shock transcription
factor"
```

```
gene id:MOB24.9"
```

```
/codon_start=1
```

```
/evidence=not experimental
```

```
/protein_id="BAB02003.1"
```

```
/db_xref="GI:9294046"
```

```
/translation="MEDDNNNNNNNNVIAPIVTKYQMVNDPSTDLITWGPANHNF
IVDPLFSQRIPLPYKHNFFNSFVRLQNTYFKVDYDPRWEFANFHELRQKHLLN
NIAARKKARGMYGODLEKIVRIERLKEORELEAIQRMNRRIEATEKRPMQMA
FLYKVVDPDLLPMLMEKERTKOQVSDKKERTVMTVKSSEERSEVERDEGRVPRV
MSSSTPFSSTENLYRHSFDGWIIVPMTCQFGSYETGLVAKSLMSNSTSTSSSLTS
TFSUPESVGGGGGGGSGIERRYKETAIFGVVSNPPPTTPYPSFLPRGG"
complement [join(37054..37590,37670..37809,37894..38197,
38288..38398,38486..38584,38662..38772,39019..39066,
39311..39367)]
/note="gene id:MOB24.10
unknown protein"
```

CDS

```
/codon_start=1
```

Alignment Scores:

Pred. No.:	7.29e-55	Length:	79706
Score:	2622.00	Matches:	573
Percent Similarity:	61.19%	Conservative:	23
Best Local Similarity:	58.83%	Mismatches:	43
Query Match:	75.93%	Indels:	337
DB:	8	Gaps:	19

US-10-086-464-2 (1-647) x AB020746 (1-79706)

Qy	1	MetSerSerAlaProSerProGlyThr-----GlySerProProSerProSer 17
Db	51787	ATGTCCACAGCGCGTCTCCAGGCACTACTCCATCACCCTCCACCGTCTCTCCACCA 51846
Qy	18	AsnSerThrThrThrProProAla-----SerAlaProProThrThrPro 35
Db	51847	AACTCGACACACCACTCTCTCCAGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 51906
Qy	36	SerSerProProPro-----ProSerThrThrThrThrThrThrThrThrThr 54
Db	51907	TCCT 51966
Qy	55	ThrProSerAlaProProProSerProProThrProThrProGlySerProPro 74
Db	51967	CCTCCT 52014

Qy	75	LeuProGlnProSerProProAlaProThrThrProGlySerProProAlaProValThr 94
Db	52015	CTTCTCTCAACTTCCCTCGCTCCCATCACTCTCTCTCTCTCTCTCTCTCTCTCTCT 52071
Qy	95	ProProThr---ArgAsnProProProSerValProGlyProProSerAsnProSerArg 113
Db	52072	CCCTCAAAACCTCGAAGCCCTCCATCTCTTAACCAAGGACCAACCAACACTCCCTCA 52128
Qy	114	GluGlyGlySerProArgProProSerSer-----ProSerProProSerProSerSer 131
Db	52129	---GGATCTACTCTAGAACCTCAACACTAAACCGTCGCGCGCTCTGATCTCTCTCC 52185
Qy	132	AspGlyLeuSerThrGlyValValValGlyLeuAlaIleGlyGlyValAlaLeuVal 151
Db	52186	GATGATGTCTACCGGAGTTGTGTAGGAATGCCATTTGGAGGAGTCGCTATTCTTGT 52245
Qy	152	IleValThrLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 169
Db	52246	ATACTGACTCTGATTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 52305
Qy	170	AspAlaTyThrValProProProProProProProProProProProProProPro 184
Db	52306	GCTGCTTACTATGTTCT 52364
Qy	184	-----
Db	52365	CAACTTCATGAATAAATAATCGAATATTAAACCATAGTTGTTCAATTCTGATCGAATTG 52424
Qy	184	-----
Db	52425	CTCAATCTGATCTCTGATTATGATTGTTATCTGTTAAAGAGATTACTTTGTGGAAAA 52484
Qy	185	-----
Db	52485	GTTCAAACTTTACAGAGATTGATGTTGTTTATCACTCGAGGACCTTACGGTGGTCAA 52544
Qy	192	GlnGlnGlnTrpArgGlnGlnAsnAlaThrProProSerAsp---HisValValThrSer 210
Db	52545	CAACAGTATTGGACAAACAAACGCGTCACGCGCGTCAGATATCATGATGATGATGATG 52604
Qy	211	LeuProProProProProProProProProProProProProProProProProPro 230
Db	52605	TTG---CCACCACCTAAGCCTCCATCTCCACCACGAAACCTCTCCGCCACCTCCACCA 52661
Qy	231	Pro---PheMetSerSerSerGlyGlySerAspTyrSerAspArgProValLeuProPro 249
Db	52662	CCAGCATTCATAGTAGTAGTAGTGGTGGTCTGACTATTTCGGATCTTCTCCCTCCA 52721
Qy	250	ProSerProGlyLeuValLeuLeuGlyPheSerLysSerThrPheThrTyrGluGluLeuAla 269
Db	52722	CCATCTCCAGGCGTGTGTAGGCTTTTCTAAAAGCACTTTCACTTAAGGAGAGTTGTG 52781
Qy	270	ArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyGlyPheGlyTyrVal 289
Db	52782	AGAGTACTAATGGCTTCTCTGAGGCTAATTTGTTAGGACAAAGGAGGTTGTTATGTG 52841
Qy	290	HisLysGlyValLeuProSerGlyLysGluValAlaValLysGlnLeuLysValGlySer 309
Db	52842	CATAAAGGTATATTGCTAGTGGGAAAGATTGCTGTGAACACAGTTGAAAGCTGGTAGT 52901
Qy	310	GlyGlnGlyGluArgGluPheGlnAlaGluValGluLeuIleSerArgValHisHisArg 329
Db	52902	GGTCAGGAGAGAGAGAGTTTCAGGCTGAGGTGAGATCATTAGCAGAGTTTCATCAGG 52961
Qy	330	HisLeuValSerLeuValGlyTyrCysIleAlaGlyAlaLysArgLeuLeuValTyrGlu 349
Db	52962	CAITTGCTTCTCTTATGTTGTTATGTTGCGCGGTGTTCAAGATTACTTGTCTATGAG 53021
Qy	350	PheValProAsnAsnAsnLeuLeuLeuHisLeuHisGlyGlu----- 363
Db	53022	TTTGTTCCAACAAACAATCTTGTGTTTTCACCTCCATCGTTA-GAACAAAACTAACTCTT 53080

QY	363	-----	363
Db	53081	TCCTTTGATTTTGTACATGCTTGGTTGGTTTTCAGTTAGTGATGATTTGTTGGGAATT	53140
QY	364	-----GlyArgProThrMetGluTrpSerThrArgLeuLysIleAla	377
Db	53141	GAATTTGTTATAGGTAAGGACGGCTACGATGGATGGATGAGTACTAGATTGAAGATTGCT	53200
QY	378	LeuGlySerAlaLysGlyLeuSerTyrLeuHisGluAspCys-----	391
Db	53201	CTTGATCTGCTAAAGGACTTTTCATATCTTCATGAAGATTCTAAAGCTCTAAACCTTAAAA	53260
QY	391	-----	391
Db	53261	ATTATGTACAGATGTTTGATAAATCTCTCTTTGTTGTCCTGCTAGGTATATATATAAAT	53320
QY	392	-----AenProLysIleIleHisArgAspIleLysAlaSerAenIleLeuIleAs	408
Db	53321	ATTTCGAGCGCAATCCGAAATCATTCACCGTGATATTAAAGGGTCAAACATATTGATTGA	53380
QY	408	pPheLysPheGluAlaLys-----	414
Db	53381	TTTCAAAATTTGAAGCTAAGGTATGTTTACTCGATATGTTTGTGAAGTGTGTTTGTAAAT	53440
QY	415	-----ValAlaAspPheGlyLeuAlaLysIleAlaAs	425
Db	53441	TGTGTCACCTAAAGCTCTCTTTCTTTTCAGGTTGCTGACTTTGGTCTTGCCGAAGATTGCTT	53500
QY	425	erAspThrAsnThrHisValSerThrArgValMetGlyThrPheGly-----	440
Db	53501	CTGTACAAACACTCATGTAICTACACCGGTGATGGGAACCTTTGGGTAAAAAATCATCTT	53560
QY	440	-----	440
Db	53561	TAATCTACTAAACATGTACATTTATGTCAGTTGCTTAAATGTCATGTCATCTTGTGCAACA	53620
QY	441	-----TyrLeuAlaProGluTyrAlaAlaSerG	450
Db	53621	AGTTCGTAACCTTTTGGTTTGGGATTTTAGGTATTTGGCTCCGGAATATGCTGCAAGTG	53680
QY	450	lyLysLeuThrGluLysSerAspValPheSerPheGlyValValLeuLeuGluLeuIleT	470
Db	53681	GAAGCTCACAGAAAAGCTGACGTTTCTCATTTGGCGTTGTACTTTTGGAACTATTGA	53740
QY	470	hrGlyArgProValAspAlaAenValTyrValAspSerLeuValAspTrp-	489
Db	53741	CTGGAGGCGCCCTGTTGATGCAACAATGCTATGTAGATGACAGCTTAGTTGACTGGG	53800
QY	489	-----	489
Db	53801	TACTTTTACTATGCTGCAGATAAATCTAAACTTCTGATTTTCTACAATTTAACTGTTGAC	53860
QY	490	-----AlaArgProLeuLeuAenArgAla	497
Db	53861	TAAACCAAAAGCTATGTGTGATTGATGTTGGTGGNACAGGCACGACCTTCTTAAACGAGCA	53920
QY	498	SerGluGlnGlyAspPheGlyLeuAlaAepAlaLysMetAsnAsnGlyTyrAspArg	517
Db	53921	TCTGAGGAAGGAGATTTTGGGGTTTGGCTGATTCAAAGATGGGTAATGAGTATGACAGA	53980
QY	518	GluGluMetAlaArgMetValAlaCysAlaAlaCysValAtcGHisSerAlaArgArg	537
Db	53981	GAGGAGATGGCTCGCATGTTGCTTGGCGTGGCGTGTGTGTCGCAATTCAGCTCGCCGC	54040
QY	538	ArgProArgMetSer-----	542
Db	54041	AGACCTCGCATGAGCGCAGGTAACCTTAATCTTCTCTAGAGAAGCATATGCTTTGG	54100
QY	542	-----	542
Db	54101	TTCTGTTTATTTGGTAGAATAATATGTTTCCGTATGTTTAAAGAAAGGTTTTCATATG	54160
QY	542	-----	542

Db	54161	GTTTTTGTGAGAAAGATCGGTTTCAGTTATGTTTATGCTTAGCCCTACTTTTTGTTTACC	54220
QY	542	-----	542
Db	54221	CATCATCATACTATCTAAACAACACCGATTGAATATGTCAAAGAAATTTGGCTGCAGTTTC	54280
QY	542	-----	542
Db	54281	TTTTGTTTTCATTTTGGCAAGGATTGTCAATTTTAAACATATGTTAATGCTAATGAGTGAG	54340
QY	543	-----GlnIleValArgAlaLeuGluGlyAsnValSerLeuSer	555
Db	54341	TGATTTGCTGTACAAAATAGCAGATAGTAGCGGGGTTAGAGGAATATGATCGCTGCT	54400
QY	556	AspLeuAenGluGlyMetArgProGlyGlnSerAenValTyrSerSerTyrGlyGlySer	575
Db	54401	GATCTTAAACGAAGGATGAGACCGGTTACACGCAACGTTATACAGCTCATATGGAGGAAGC	54460
QY	576	ThrAspTyrAspSerSerGlnTyrAsnGluAspMetLysLysPheArgLysMetAlaLeu	595
Db	54461	ACAGACTATGACACGACGCAATACACGACGACATGATAAAGTTTAGGAAATGGCTCTT	54520
QY	596	GlyThrGlnGluTyrAsnAlaThrGlyGluTyrSerAenProThrSerAspTyrGlyLeu	615
Db	54521	GGAACTCAAGNATACGCCACACCGCGGAGTACAGTAATCCACAGTACTCGGACTG	54580
QY	616	TyrProSerGlySerSerSerGluGlyGlnThrThrArgGluMetGluMetGlyLysIle	635
Db	54581	TACCGGCTGGTTCAAGCAGTGAAGGTCAAGCCACACGAGAAATGGAGATGGAAAGATT	54640
QY	636	LysArgThrGlyGlnGlyTyrSerGlyProSerLeu	647
Db	54641	AAGAAACCCGTCAGGTATAGTGGACCCCTCTCTT	54676

RESULT 15	AKI03247	2307 bp	mRNA	linear	PLN 24-JUL-2003
LOCUS	AKI03247				
DEFINITION	Oryza sativa (japonica cultivar-group) cDNA clone:J033123K23, full insert sequence.				
ACCESSION	AKI03247				
VERSION	AKI03247.1	GI:32988456			
KEYWORDS	FLI cDNA; CAP trapper.				
SOURCE	Oryza sativa (japonica cultivar-group)				
ORGANISM	Oryza sativa (japonica cultivar-group)				
REFERENCE	1				
AUTHORS	The Rice Full-Length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-Length cDNA Project Team; Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group; Otono, Y., Murakami, K., Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y., Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M., Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J., Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J., Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN; Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Oka, Y., Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Yoshino, M. and Hayashizaki, Y.				
TITLE	Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice				
JOURNAL	Science 301 (5631), 376-379 (2003)				
MEDLINE	22752273				
PUBMED	12869764				
REFERENCE	2	(bases 1 to 2307)			
AUTHORS	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K.,				

Job time : 5986 secs

```
Db 988 |||||TATGTTCAAGAGGATGTCGCGAATGGGACAGAGGTGCTGTGAACAATGAGAGAT 1047
Qy 308 GlySerGlyGlnGluArgGluPheGlnAlaGluValGluIleIleSerArgValHis 327
Db 1048 GGGAGTGGGACGGAGAGCGTGATTTTCAGGCGGAGGTTGAGATTATCAGCGGGTACAT 1107
Qy 328 HisArgHisLeuValSerLeuValGlyTyrCysIleAlaGlyAlaLysArgLeuVal 347
Db 1108 CACAAGCATCTTGTAACATGGTTGGTTATTGCAATTCCTGGAGGAGAGGTTGCTTGTG 1167
Qy 348 TyrGluPheValProAsnAsnLeuGluLeuHisLeuHisGlyGluGlyArgProThr 367
Db 1168 TATGAGTATGTTCCAAAACAACATTCAGGCTCCACTTGCATGGGAGAGCGGACCAACC 1227
Qy 368 MetGluTrpSerThrArgLeuLysIleAlaLeuGlySerAlaLysGlyLeuSerTyrLeu 387
Db 1228 ATGGNATGGCTTACAAAGATTACGCATTGCTTTGGGTGCTGCGAAGGGATTGGCATACTT 1287
Qy 388 HisGluAspCysAsnProLysIleIleHisArgAspIleLysAlaSerAsnIleLeu 407
Db 1288 CATGAAGATTGCCATCTCAAGATCATTCACCGTGACATAAAAGTCAGCAAAATATTCTTCT 1347
Qy 408 AspPheLysPheGluAlaLysValAlaAspPheGlyLeuAlaLysIleAlaSerAspThr 427
Db 1348 GACGCGAGATTGAGGCAAGAGTGCGAGATTTCGGACTTCGAAATGACTTCTGATAAT 1407
Qy 428 AsnThrHisValSerThrArgValMetGlyThrPheGlyTyrLeuAlaProGluTyrAla 447
Db 1408 AACACTCATGTTTCAACACAGAGTAATGGGCACATTTGGGTACCTTGCACAGAGTATGCG 1467
Qy 448 AlaSerGlyLysLeuThrGluLysSerAspValPheSerPheGlyValValLeuLeuGlu 467
Db 1468 TCTTCTGGCCAGCTAACTGAGAAATCAGATGCTTTCTTTCGGAGTAATGCTTCTTGAG 1527
Qy 468 LeuIleThrGlyArgProValAspAlaAsnAsnValTyrValAspAspSerLeuVal 487
Db 1528 CTAATAACTGGCCGCGCTTAAATGATGCGAGCATCAGATGATGCAATTATGATGCTTAGTG 1587
Qy 488 AspTrpAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGluGlyLeuAla 507
Db 1588 GATTGGGCAAGGCTTTAATGATGCGAGCATCAGATGATGCAATTATGATGCTTAGTG 1647
Qy 508 AspAlaLysMetAsnAsnGlyTyrAspArgGluGluMetAlaArgMetValAlaCysAla 527
Db 1648 GATCACGCGTAGGACAGAGTACAAATGTTAAACGAGATGGCGAGGATGATTGCTTGTGCA 1707
Qy 528 AlaAlaCysValArgHisSerAlaArgArgProArgMetSerGlnIleValArgAla 547
Db 1708 GCTGCTTGTGTACGCCATTCTGCACGCGCGCCGACCGGATGAGCGGTTGTTGCGGCC 1767
Qy 548 LeuGluGlyAsnValSerLeuSerAspLeuAsnGluGlyMetArgProGlyGlnSerAsn 567
Db 1768 TTGGAAGCGGATGTCATCGGATGATCTGAACGAAGGTGTTGCTCTGGCCATAGCCGA 1827
Qy 568 ValTyrSerSerTyrGlyGlySerThrAspTyrAspSerSerGlnTyrAsnGluAspMet 587
Db 1828 TTCTTGGGATCATAC---AACAGCAACAGTACGACACGGGCCATTACAATGAAGACCTG 1884
Qy 588 LysLysPheArgLysMetAlaLeuGlyThrGlnGluTyrAsnAlaThrGlyGluTyrSer 607
Db 1895 AAGAAATTCAAGAAATGGCATTGGCAGTGGCAATCTA-----GAGAGCAGC 1932
Qy 608 AsnProThrSerAspTyrGlyLeuTyrProSerGlySerSerSerGluGlyGlnThrThr 627
Db 1933 CAGCAACCGCAGCCACCGAGTTCGTCGGAATCGATCCGATATCAATGGGGATGCGCGT 1992
Qy 628 Arg-----GluMetGluMetGlyLysIleLysArgThrGly 639
Db 1993 CAAATACCGGACGAGATGGAGATGGGAGTTTGAAGAAAGATGCG 2040
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus p2n model

Run on: May 12, 2004, 20:46:26 ; Search time 603 Seconds
(without alignments)
4558.183 Million cell updates/sec

Title: US-10-086-464-2
Perfect score: 3453
Sequence: 1 MSSAPSGTGTGPPPPSNT.....REMGKIKRTGQYSGPSL 647

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgn2.1/USPTO_spool_p/US10086464/runat_11052004_121814_7833/app query.fasta_1.839
-DB=N Geneseq 29Jan04 -QMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bites -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORF=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10086464.ecgn.1.1.708 -runat_11052004_121814_7833 -NCPU=6 -ICPU=3
-NO MMAP -LARGESQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FCGAPOP=6
-FCGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq 29Jan04: *
1: Geneseqn1980s: *
2: Geneseqn1990s: *
3: Geneseqn2000s: *
4: Geneseqn2001as: *
5: Geneseqn2001bs: *
6: Geneseqn2002s: *
7: Geneseqn2003as: *
8: Geneseqn2003bs: *
9: Geneseqn2003cs: *
10: Geneseqn2004s: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3453	100.0	1944	4	Aaf77094 Brassica
2	3453	100.0	1944	9	Acf36548 B. napus
3	3453	100.0	2189	9	Acf36549 B. napus
4	2936.5	85.0	1959	9	Acf36557 A. thalia
5	2885	83.6	2324	3	Aac36968 Arabidops
6	1821.5	52.8	1812	9	Acf36554 O. sativa
7	1668	48.3	1515	9	Acf36555 Z. mays P
8	1637.5	47.4	1488	9	Acf36556 G. max PE

9	1507	43.6	2520	4	Aaf77096	Arabidops
10	1504	43.6	2520	9	ACF36551	Arabidops
11	1479.5	42.8	2820	4	Aaf77097	Arabidops
12	1471.5	42.6	2820	9	ACF36552	Arabidops
13	1209.5	35.0	2880	4	Aaf77095	Arabidops
14	1209.5	35.0	2880	9	ACF36550	Arabidops
15	1185	34.3	3060	9	ACF36553	Arabidops
16	1181	34.2	3060	4	Aaf77098	Arabidops
17	839.5	24.3	1923	7	ADA70890	Rice gene
18	816.5	23.6	2124	6	ABZ13274	Arabidops
19	804.5	23.3	2865	7	ADA70701	Rice gene
20	802.5	23.2	2011	4	AAH22557	P. patens
21	797	23.1	1730	7	ADA70647	Rice gene
22	793.5	23.0	2823	7	ADA69478	Rice gene
23	782.5	22.7	1419	3	AAC43029	Arabidops
24	782.5	22.7	1419	6	ABZ13121	Arabidops
25	775.5	22.5	1626	3	AAC48313	Arabidops
26	775.5	22.5	1917	7	ADA70527	Rice gene
27	773	22.4	1479	3	AAC43040	Arabidops
28	769.5	22.3	1353	3	AAC42860	Arabidops
29	769.5	22.3	1353	6	ABZ14106	Arabidops
30	766	22.2	1494	3	AAC48560	Arabidops
31	758.5	22.0	1572	3	AAC42898	Arabidops
32	758.5	22.0	1572	6	ABZ13705	Arabidops
33	758.5	22.0	1920	6	AAL45905	A thalian
34	758.5	22.0	2012	3	AAC44847	Arabidops
35	756	21.9	3239	3	AAC79271	Eucalyptu
36	754.5	21.9	2136	7	ADA69829	Rice gene
37	749.5	21.7	1751	3	AAC45313	Arabidops
38	744.5	21.6	1485	3	AAC42702	Arabidops
39	740.5	21.4	3525	7	ADA71170	Rice gene
40	740	21.4	2868	3	AAA79266	Pinus rad
41	740	21.4	3153	7	ADA71066	Rice gene
42	736.5	21.3	1703	3	AAC46629	Zea mays
43	736	21.3	3624	7	ADA69512	Rice gene
44	735	21.3	2004	3	AAC40780	Arabidops
45	734	21.3	1455	3	AAC45877	Arabidops

ALIGNMENTS

RESULT 1
Aaf77094
ID Aaf77094 standard; DNA; 1944 BP.
XX
AC Aaf77094;
XX
DT 17-MAY-2001 (first entry)
XX
DE
DE Brassica napus PERK1 DNA.
XX
KW Proline-rich extensin-like receptor kinase; PERK; resistance; plant; ds.
XX
OS Brassica napus.
XX
PN WO200114563-A1.
XX
PD 01-MAR-2001.
XX
PF 18-AUG-2000; 2000WO-CA000966.
XX
PR 19-AUG-1999; 99US-0149466P.
PR 13-OCT-1999; 99US-0159122P.
XX
(GORI/) GORING D.
(SILV/) SILVA N.
PA
PA
XX
XX Goring D, Silva N;
XX
DR WPI; 2001-244305/25.
XX
XX New proline-rich, extensin-like receptor kinase nucleic acids and
PT polypeptides useful for increasing plant wounding or pathogen resistance,

PT or for producing transgenic plants with increased wounding or pathogen
 PT resistance.

PS Claim 6; Fig 1; 91pp; English.

XX The present invention relates to proline-rich extensin-like receptor
 CC kinase (PERK). The PERK nucleic acids and polypeptides are useful for
 CC increasing the resistance of plants to wounding and pathogens. These are
 CC also useful for producing transgenic plants with increased wounding and
 CC pathogen resistance compared with a wild type plant, as well as in assays
 CC for identifying and developing compounds to inhibit and/or enhance
 CC polypeptide function directly

XX Sequence 1944 BP; 464 A; 550 C; 459 G; 471 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 4, 62e-93 Length: 1944
 Score: 3453.00 Matches: 647
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 4 Gaps: 0

US-10-086-464-2 (1-647) x AAF77094 (1-1944)

Qy 1 MetSerSerAlaProSerProGlyThrGlySerProProSerProProSerAsnSerThr 20
 Db 1 ATGTCCTCGGCGCGTCTCCGGGACTGGTTCGCCCTCCATCTCCACCATCAAACTCCACA 60
 Qy 21 ThrThrThrProProAlaSerAlaProProProThrThrProSerSerProPro 40
 Db 61 ACCACCATCTCTCCAGCTTCGGCTCTCTCCACACACACCTTCTCTCTCCGCGG 120
 Qy 41 ProSerThrIleProThrSerProProSerSerArgSerThrProSerAlaProPro 60
 Db 121 CCATCCATATTCGAGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180
 Qy 61 ProSerProProThrProSerThrProGlySerProProProProLeuProGlnProSerPro 80
 Db 181 CCATCTCCACCAACTCCATCTACGGCGGATCTCCACCTCTCTCTCTCTCTCTCTCTCTCA 240
 Qy 81 ProAlaProThrThrProGlySerProProAlaProValThrProProThrArgAsnPro 100
 Db 241 CCCGCTTCAACTACCGCCGGATCTCCACCGGACCTGTTACTCTCTCTCTCTCTCTCTCTCT 300
 Qy 101 ProProSerValProGlyProProSerAsnProSerArgGlyGlySerProArgPro 120
 Db 301 CCACCTTCAGTCCAGAGACCCGCTCCAACTCTTCCACCGAAGGAGATCTCTCGACCT 360
 Qy 121 ProSerSerProSerProSerProSerSerSerAspGlyLeuSerThrGlyValVal 140
 Db 361 CCATCTTCTCTCCGCGCGTCT 420
 Qy 141 GlyIleAlaIleGlyGlyValAlaLeuLeuValIleValThrLeuIleCysLeuLeuCys 160
 Db 421 GGAATCGGCATCGAGGAGTCTCTCTCTGTGTGAGTAGTACTCTGATTGTCTCTCTCTGT 480
 Qy 161 LysLysLysArgArgArgAspGluAspAlaTyrTyrValProProProProProPro 180
 Db 481 AAGAGAAGACGCGAG 540
 Qy 181 GlyProLysAlaGlyGlyProTyrGlyGlyGlnGlnGlnTTPArgGlnGlnAsnAla 200
 Db 541 GGTCCCAAGCGGAGGACCTTACCGGTGGACAGCAGCAACAATGCGGCGGCAACAAACGCA 600
 Qy 201 ThrProProSerAspHisValValThrSerLeuProProProProProProProPro 220
 Db 601 ACACACCGCTCAGATCATGTGTGACGTACACTTACCACCAACCACTTAAGGCTCCATCTCCA 660
 Qy 221 ProArgGlnProProProProProProProProPheMetSerSerSerGlyGlySerAsp 240
 Db 661 CCACGGCAACCTCTCTCCACCTCCACCGCGCTTTCATGAGCAGCAGCGCGGCTCCGAC 720

Qy 241 TyrSerAspArgProValLeuProProSerProGlyLeuValLeuGlyPheSerIys 260
 Db 721 TACTCGACCGTCCAGTTCTTCTCCACCGTCTCCAGGGCTGTGTAGGCTTCTCCAAA 780
 Qy 261 SerThrPheThrTyrGluGluLeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeu 280
 Db 781 AGCATTTTCATACAGAGGAGCTAGCTAGAGCCCACTATGGTTTCTCCGAGCGCACTTG 840
 Qy 281 LeuGlyGlnGlyGlyPheGlyTyrValHisLysGlyValLeuProSerGlyLysGluVal 300
 Db 841 TTAGACACAGCGGTTCCGTTACGTGCACAAAGTGTGTTCCTAGTGGGAAAGAGTT 900
 Qy 301 AlaValLysGlnLeuLysValGlySerGlyGlnGlyGluArgGluPheGlnAlaGluVal 320
 Db 901 GCTGTGAAGCAGTTGAAGTTGGAGTGGTCAGGAGAGAGGAGTTTCAGCGAGAGGTT 960
 Qy 321 GluIleIleSerArgValHisHisArgHisLeuValSerLeuValGlyTyrCysIleAla 340
 Db 961 GAGATCATCAGCAGATTCCACACAGCACTGGTGTCTCTTGTGGTATTATGATCGCC 1020
 Qy 341 GlyAlaLysArgLeuLeuValTyrGluPheValProAsnAsnLeuGluLeuHisLeu 360
 Db 1021 GGTGCCAAAGATTGCTTCTATGAGTTTGTCTTAACAACAATCTCGAGCTTCACCTC 1080
 Qy 361 HisGlyGluGlyArgProThrMetGluThrSerThrArgLeuLysIleAlaLeuGlySer 380
 Db 1081 CATCGCAGGAGCGCTTACAAATGGAATGGAGCACCAGATTGAAGATTGCTTCTGGATCT 1140
 Qy 381 AlaLysGlyLeuSerTyrLeuHisGluAspCysAsnProLysIleIleHisArgAspIle 400
 Db 1141 GCTAAGAGCTTCTTATCTTCATGAAGTTGCAATCTTAANAATCATTCACCGTGATC 1200
 Qy 401 LysAlaSerAsnIleLeuIleAspPheLysPheGluAlaLysValAlaAspPheGlyLeu 420
 Db 1201 AAGGCTTCAACAATATTGATAGATTTCAGTTTGAAGCTAAGTTGCTGTGATTTGGTCTT 1260
 Qy 421 AlaLysIleAlaSerAspThrAsnThrHisValSerThrArgValMetGlyThrPheGly 440
 Db 1261 GCTAAGATTGCTTCTGATACAAACACGATGTATCAACACGCTGTGATGGGAACCTTTGGG 1320
 Qy 441 TyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPheSer 460
 Db 1321 TACTTGGCTCGGAAATACGCTGCAAGCGGAAGCTCACGGAGAAGTCTGACGTTTCTCA 1380
 Qy 461 PheGlyValValLeuLeuGluLeuIleThrGlyArgArgProValAspAlaAsnVal 480
 Db 1381 TTTGGCGTTGTCTTTTGGAGCTCATTTACTTGAGCTCGACCTCGACCTTGATGCCAACAAATGTC 1440
 Qy 481 TyrValAspSerLeuValAspTyrAlaArgProLeuLeuAsnArgAlaSerGluGln 500
 Db 1441 TATGTAGATGACAGCTTAGTTGACTGGGCGACGACCATTTGCTTAAACCGAGCATCTGAGCA 1500
 Qy 501 GlyAspPheGluGlyLeuAlaAspAlaLysMetAsnAsnGlyTyrAspArgGluGluMet 520
 Db 1501 GGAGACTTGGAGGTTAGCTGATGCAAGATGAATAATGGGTATGACAGAGAGAGATG 1560
 Qy 521 AlaArgMetValAlaCysAlaAlaAlaCysValArgHisSerAlaArgArgProArg 540
 Db 1561 GCTCGCATGGTTGCTTGTCTCGGCTTGTTCGCCATTTCAGCTCGCGCGCAGACCTCGC 1620
 Qy 541 MetSerGlnIleValArgAlaLeuGluGlyAsnValSerLeuSerAspLeuAsnGluGly 560
 Db 1621 ATGAGCCAGATTGTCGTCGCTTAGAAGAAATGTATCCTGTGTGATCTTAAACGAAAGG 1680
 Qy 561 MetArgProGlyGlnSerAsnValTyrSerSerTyrGlyGlySerThrAspTyrAspSer 580
 Db 1681 ATGAGACCAAGTCAAGCAATGTATACAGCTCATACGGAGAGAGACCGATTATGACTCG 1740
 Qy 581 SerGlnTyrAsnGluAspMetLysPhePheArgLysMetAlaLeuGlyThrGlnGluTyr 600
 Db 1741 AGCCAGTACAAATGAAGACATGAAGAAGTTTAGGAAAAATGGCACTTGGAACTCAAGAGTAC 1800
 Qy 601 AsnAlaThrGlyGluTyrSerAsnProThrSerAspTyrGlyLeuTyrProSerGlySer 620

Db 1801 AACGCCAGGGTGTAGTACAGTAATCCGACCACTGACTATGAGCTGTACCCGCTGTGTTCA 1860
Qy 621 SerSerGluGlyGlnThrThrArgGluMetGluMetGlyLysIleLysArgThrGlyGln 640
Db 1861 AGCAGCGAGGCGCCAAACACACCGGAATGAGATGGGGAAGATTAAAGAGAACCGGTGAG 1920
Qy 641 GlyTyrSerGlyProSerLeu 647
Db 1921 GGTATAGTGACCTTCTCTT 1941

RESULT 2

ACF36548
ID ACF36548 standard; cDNA; 1944 BP.

XX AC ACF36548;

DT 18-DEC-2003 (first entry)

XX B. napus PERK1 receptor kinase encoding cDNA.

DE PERK1; transgenic; plant; proline-rich extensin-like receptor kinase;
KW wound; pathogen resistance; plant growth; seed production; gene; ss.

XX Brassica napus.

XX Key Location/Qualifiers

FT CDS 1..1944

FT /*tag= a

FT /transl_except= (pos: 1332..1334, aa: Phe)

FT /product= "PERK1"

FT /note= "Proline-rich Extensin-like Receptor Kinase"

PN W02003072763-A1.

XX 04-SEP-2003.

XX 28-FEB-2003; 2003WO-CA000274.

XX 28-FEB-2002; 2002CA-02373903.

PR 28-FEB-2002; 2002US-00086464.

XX (GORI/) GORING D.

PA (SILV/) SILVA N.

PA (HAFF/) HAFFANI Y Z.

XX Goring D, Silva N, Haffani YZ;

XX WPI; 2003-712727/67.

DR P-PSDB; ABR82937.

XX Producing a transgenic plant having an increased plant resistance, plant growth or seed production comprises transforming a plant with a nucleic acid molecule having a Proline-rich Extensin-like Receptor Kinase activity.

XX Claim 3; Fig 1A; 123pp; English.

XX The invention relates to producing a transgenic plant having increased plant height, number of branches, number of seed pods and/or seed production compared to a non-transgenic plant, and/or quicker flowering or later senescence compared to a non-transgenic plant. The method involves transforming a plant with a vector including a Proline-rich Extensin-like Receptor Kinase (PERK) nucleic acid molecule or a nucleic acid molecule having PERK activity. The method, as well as the PERK nucleic acid molecule and polypeptide, are useful in increasing plant resistance to wounding and pathogens and in increasing plant growth and seed production. The nucleic acid molecule and polypeptide may also be used in producing transgenic plants or transgenic host cells. The present sequence represents a cDNA encoding a B. napus PERK1 receptor kinase polypeptide

XX Sequence 1944 BP; 464 A; 550 C; 459 G; 471 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 4.62e-93 Length: 1944
Score: 3453.00 Matches: 647
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-10-086-464-2 (1-647) x ACF36548 (1-1944)

Qy 1 MetSerSerAlaProSerProGlyThrGlySerProProSerProProSerSerSerThr 20
Db 1 ATGTCTCGCGCGGTCTCGGGGACTGGTTCGCTCCATCTCCACCATCAAACTCCACA 60
Qy 21 ThrThrThrProProAlaSerAlaProProThrThrProSerSerProPro 40
Db 61 ACCACCACTCTCTCTCAGCTTCGCTCCCTCCACCAACACCTTCTTCTCTCCGCG 120
Qy 41 ProSerThrThrProProProProSerSerSerArgSerThrProSerAlaProPro 60
Db 121 CCATCCACTATTCGACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180
Qy 61 ProSerProProThrProSerThrProGlySerProProProProGlnProSerPro 80
Db 181 CCATCTCCACCACTCCATCTAGCGCGGATCTCCACCTCTCTCTCTCTCTCTCTCT 240
Qy 81 ProAlaProThrThrProGlySerProProAlaProValThrProThrArgAsnPro 100
Db 241 CCGGCTCAACTAGCGCGGATCTCCACCGCGACTCTTACTCTCTCTCTCTCTCTCT 300
Qy 101 ProProSerValProGlyProProSerAsnProSerArgGluGlyGlySerProArgPro 120
Db 301 CCACCTTCAGTCCAGGACCACTCCATCTCTCCACCGGTTTATCAACAGGAGTGGTG 360
Qy 121 ProSerSerProProProProSerProSerSerAspGlyLeuSerThrGlyValVal 140
Db 361 CCATCTCTCTCGCGCGGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 420
Qy 141 GlyIleAlaIleGlyGlyValAlaLeuLeuValIleValThrLeuIleCyeLeuLeu 160
Db 421 GGAATGCCATCGAGGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 480
Qy 161 LysLysLysArgArgArgAspGluAspAlaTyrTyrValProProProProProPro 180
Db 481 AAGAAGAAACGACGAGAGACGAGAGATGCTTACTATGTTCTCTCGCCACCTCTCT 540
Qy 181 GlyProLysAlaGlyGlyProTyrGlyGlyGlnGlnGlnGlnTrpArgGlnAsnAla 200
Db 541 GGTCCCAAGCGGAGGACCTTACGGTGGACAGCAGCAACAATGGCGGCAACAAACGCA 600
Qy 201 ThrProProSerAspHisValValThrSerLeuProProProProProProProPro 220
Db 601 ACACCACTCGATATGTCGAGCTACTACCACTACCACTACCACTAAGGCTCCATCTCA 660
Qy 221 ProArgGlnProProProProProProPheMetSerSerSerSerGlyGlySerAsp 240
Db 661 CCAGGCAACTCTCTCCACTCCACCTCCACCTCCATGAGCAGCAGCGGGCTCCGAC 720
Qy 241 TyrSerAspArgProValLeuProProProProGlyLeuValLeuGlyPheSerLys 260
Db 721 TACTCGACCGCTCAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 780
Qy 261 SerThrPheThrTyrGluLeuAlaAlaArgAlaThrAsnGlyPheSerGlyAlaAsnLeu 280
Db 781 AGCACTTTCATACGAGGAGCTAGCTAGAGCCACCAATGGTTTCTCCGAGCGCACTTG 840
Qy 281 LeuGlyGlnGlyGlyPheGlyTyrValHisLysGlyValLeuProSerGlyLysGluVal 300
Db 841 TTAGGACNAGCGGGTTCGGTTACGTGCACAAAGGTGTGTCTAGTGGGAAGAGTT 900
Qy 301 AlaValLysGlnLeuLysValGlySerGlyGlnGlyGluArgGluPheGlnAlaGluVal 320

Db 901 GCTGTGAAGACAGTTCAAAAGTTGGAGTGGTTCAGGAGAGAGGAGTTTCAGGCAGAGGTT 960
Qy 321 GluLeuIleSerArgValHisArgHisLeuValSerLeuValGlyTyrCysIleAla 340
Db 961 GAGATCATCAGCAGAGATTCCACACAGGCATCTGGTGTCTCTTGTGGTTATTGTCATCGCC 1020
Qy 341 GlyAlaLysArgLeuLeuValTyrGluPheValProAsnAsnAsnLeuGluLeuHisLeu 360
Db 1021 GGTGCCAAAAGATTGCTTGTCTATGAGTTTGTCTTAAACAATCTCGAGCTTCACCTC 1080
Qy 361 HisGlyGluGlyArgProThrMetGluTrpSerThrArgLeuLysIleAlaLeuGlySer 380
Db 1081 CATGGCGAGGACGGCCTCAATGAATGGAGCACCAGATTGAAGATTGCTCTTGGATCT 1140
Qy 381 AlalysGlyLeuSerTyrLeuHisGluAspCysAsnProLysIleIleHisArgAspIle 400
Db 1141 GCTAAAGGACTTCTCTATCTTCATGAAGATTGCAATCTTAAATCAATTCACCGTGATATC 1200
Qy 401 LysAlaSerAsnIleLeuIleAspPheLysPheGluAlaLysValAlaAspPheGlyLeu 420
Db 1201 AAGGCTTCAAAACATATTGATAGATTCAAGTTTGAAGCTAAGGTTGCTGATTTTGGTCTT 1260
Qy 421 AlalysIleAlaSerAspThrAsnThrHisValSerThrArgValMetGlyThrPheGly 440
Db 1261 GCTAAGATTGCTCTGATACAAACGCGATGATCAACACGCTGTGATGGGAACCTTTGGG 1320
Qy 441 TyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPheSer 460
Db 1321 TACTTGGCTCGGAATACGCTCAAGCGGAAGCTCACGGAAGTCTGACGTTTCTCA 1380
Qy 461 PheGlyValValLeuLeuGluLeuIleThrGlyArgArgProValAspAlaAsnVal 480
Db 1381 TTTGGCGTGTGCTTTTGGAGCTATTACTGAGCTGACGCCGTTGATGCCCAACAATGTC 1440
Qy 481 TyrValAspAspSerLeuValAspTrpAlaArgProLeuLeuAsnArgAlaSerGluGln 500
Db 1441 TATGTAGATGACAGCTTAGTTGACTGGCGACGACCAATGCTTTAACCGAGCATCTGAGCAA 1500
Qy 501 GlyAspPheGluGlyLeuAlaAspAlaLysMetAsnGlyTyrAspArgGluMet 520
Db 1501 GGAGACTTTGAGGGTTAGCTGATGCCAAAGATGAATAATGGGTATGCACAGAGAGAGATG 1560
Qy 521 AlaArgMetValAlaCysAlaAlaCysValArgHisSerAlaAtqArgArgProArg 540
Db 1561 GCTGCATGGTGTGCTGTGCTGCGGCTTGTGTGCCATTGAGCTCGCCGAGACCTCGC 1620
Qy 541 MetSerGlnIleValArgAlaLeuGluGlyAsnValSerLeuSerAspLeuAsnGluGly 560
Db 1621 ATGAGCCAGATTGTCGTGCGTTAGAGAGAAATGTATCACTGCAGATCTTAAACGAAGG 1680
Qy 561 MetArgProGlyGlnSerAsnValTyrSerSerTyrGlyGlySerThrAspTyrAspSer 580
Db 1681 ATGAGACCAGGTCAAAAGCAATGTATACAGCTCATACGGAGGAAGCACCGATTATGACTCG 1740
Qy 581 SerGlnTyrAsnGluAspMetLysLysPheArgLysMetAlaLeuGlyThrGlnGluTyr 600
Db 1741 AGCCAGTACAATGAAGACATGAAGAAAGTTTAGGAAAATGGCACCTGGAACTCAGAGTAC 1800
Qy 601 AsnAlaThrGlyGluTyrSerAsnProThrSerAspTyrGlyLeuTyrProSerGlySer 620
Db 1801 AACGCCACGGGTGAGTACAGTAATCCGACCATGACTATGACTGTACCCGCTCGTTTCA 1860
Qy 621 SerSerGluGlyGlnThrThrArgGluMetGluMetGlyLysIleLysArgThrGlyGln 640
Db 1861 AGCAGCAGGGCGCAACACACGCAATAGAGATGGGGAAGATTAAAGAAACCGGTGAC 1920
Qy 641 GlyTyrSerGlyProSerLeu 647
Db 1921 GGTATAGTGGACCTTCTCTT 1941

RESULT 3
ACF36549
ID ACF36549 standard; cDNA; 2189 BP.

XX ACF36549;
AC 18-DEC-2003 (first entry)
XX DE B. napus PERK1 receptor kinase encoding cDNA.
XX PERK1; transgenic; plant; proline-rich extensin-like receptor kinase;
KW wound; pathogen resistance; plant growth; seed production; gene; ss.
XX Brassica napus.
XX Key Location/Qualifiers
CDS 1..2189
FT /*tag= C
FT /note= "ABR82938; this protein contains the amino acids
FT corresponding to 5' and 3'UTR regions, though only the
FT relevant aa residues (ABR82937) is used in the invention"
FT 1..96
FT /*tag= a
FT 97..2040
FT /*tag= b
FT /transl_except= (pos: 1332..1334, aa: Phe)
FT /product= "PERK1"
FT /note= "ABR82937; proline-rich Extensin-like Receptor
FT Kinase"
FT 3'UTR 2041..2189
FT /*tag= d
XX WO2003072763-A1.
XX 04-SEP-2003.
XX 28-FEB-2003; 2003WO-CA000274.
XX 28-FEB-2002; 2002CA-02373903.
XX 28-FEB-2002; 2002US-00086464.
XX (GORI/) GORING D.
XX (SILV/) SILVA N.
XX (HAFF/) HAFFANI Y Z.
XX Goring D, Silva N, Haffani YZ;
XX WPI; 2003-712727/67.
XX P-ESDB; ABR82937, ABR82938.
XX Producing a transgenic plant having an increased plant resistance, plant
XX growth or seed production comprises transforming a plant with a nucleic
XX acid molecule having a Proline-rich Extensin-like Receptor Kinase
XX activity.
XX Disclosure; Fig 1D; 123pp; English.
XX The invention relates to producing a transgenic plant having increased
XX plant height, number of branches, number of seed pods and/or seed
XX production compared to a non-transgenic plant, and/or quicker flowering
XX or later senescence compared to a non-transgenic plant. The method
XX involves transforming a plant with a vector including a proline-rich
XX Extensin-like Receptor Kinase (PERK) nucleic acid molecule or a nucleic
XX acid molecule having PERK activity. The method, as well as the PERK
XX nucleic acid molecule and polypeptide, are useful in increasing plant
XX resistance to wounding and pathogens, and in increasing plant growth and
XX seed production. The nucleic acid molecule and polypeptide may also be
XX used in producing transgenic plants or transgenic host cells. The present
XX sequence represents a cDNA encoding a B. napus PERK1 receptor kinase
XX polypeptide
SQ Sequence 2189 BP; 544 A; 592 C; 494 G; 559 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 5e-93 Length: 2189
Score: 3453.00 Matches: 647

Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-10-086-464-2 (1-647) x ACF36549 (1-2189)

Qy	1	MetSerSerAlaProSerProGlyThrGlySerProProSerProProSerAsnSerThr	20
Db	97	ATGTCCTCGGCGCGCTCTCGGGGACTGGTTGGCTCTCCATCTCCACCATCAAACTCCACA	156
Qy	21	ThrThrThrProProProAlaSerAlaProProProThrThrProProSerProProPro	40
Db	157	ACCACCACTCTCTCCAGTTCGGCTCTCCCTCCACACACACCTCTCTCTCTCTCCGCGG	216
Qy	41	ProSerThrThrProThrSerProProProSerSerArgSerThrProSerAlaProPro	60
Db	217	CCATCCACTATTCGACATCT	276
Qy	61	ProSerProProThrProSerThrProGlySerProProProProProProProProPro	80
Db	277	CCATCTCCACCACTCCATCTACGCGGGATCTCCACCTCTCTCTCTCTCTCTCTCTCTCA	336
Qy	81	ProAlaProThrThrProGlySerProProAlaProValThrProProThrArgAsnPro	100
Db	337	CCCGCTCCAACTAGCGCCGGATCTCCACCCGACCTGTCTCTCTCTCTCTCTCTCTCTCT	396
Qy	101	ProProSerValProGlyProProProSerAsnProSerArgGlyGlySerProArgPro	120
Db	397	CCACCTTCAGTCCCGAGGACCCGCTCAATCTCTCACCGGAGGAGATCTCTCTCGACCT	456
Qy	121	ProSerSerProProProProSerSerSerAspGlyLeuSerThrGlyValValVal	140
Db	457	CCATCTCTCTCTCTCGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	516
Qy	141	GlyThrAlaThrGlyValAlaLeuLeuValThrValThrLeuLeuLeuLeuLeuLeuCys	160
Db	517	GGAACTCGCATCGGAGGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	576
Qy	161	LysLysLysArgArgArgAspGluGluAspAlaTyrValProProProProProProPro	180
Db	577	AAGAAGAACACCGAGAGAGAGAGAGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	636
Qy	181	GlyProLysAlaGlyGlyProTyrGlyGlyGlnGlnGlnGlnGlnGlnGlnGlnGlnGln	200
Db	637	GGTCCCAAGCGGAGGACCTTACGGTGGACAGAGACAACTAGCGGCAACAAACGCA	696
Qy	201	ThrProProSerAspHisValValThrSerLeuProProProProProProProProPro	220
Db	697	ACACCACCGTCAGATCATGTCTGAGTCACTACCCACCACTAAGGCTCTCATCTCCA	756
Qy	221	ProArgGlnProProProProProProProProProProProProProProProProPro	240
Db	757	CCACGGCAACCTCTCCACCTCCACCCGCTTTTCATGACGACGCGCGGCTCCGAC	816
Qy	241	TyrSerAspArgProValLeuProProProProProProProProProProProProPro	260
Db	817	TACTCGGACCGTCCAGTCTCTCTCTCCACCGTCTCCAGGCTTGTAGGCTTCTCCAA	876
Qy	261	SerThrPheThrTyrGluGluLeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeu	280
Db	877	AGCACTTTCATACGAGGAGCTAGTAGGCCCACTAATGGTTCTCCGAGGCGCACTTG	936
Qy	281	LeuGlyGlnGlyGlyPheGlyTyrValHisLysGlyValLeuProProProProProPro	300
Db	937	TTAGGACAAAGCGGGTTCGGTTCAGTCCACAAAGGTGTGTTCCTCTCTCTCTCTCTCT	996
Qy	301	AlaValLysGlnLeuLysValGlySerGlyGlnGlyGluArgGluPheGlnAlaGluVal	320
Db	997	GCTGTGAACCAAGTTGAAGTTGGAGTGTCTCAGGAGAGAGAGGATTTTCAGGCGAGG	1056
Qy	321	GluThrLeuSerArgValHisArgHisLeuValSerLeuValGlyTyrCysLeuAla	340

RESULT 4

ACF36557 ID ACF36557 standard; DNA; 1959 BP.

XX ACF36557;

XX DT 18-DEC-2003 (first entry)

XX A. thaliana PERK1 protein encoding genomic DNA.

XX PERK1; transgenic; plant; proline-rich extensin-like receptor kinase;

KW wound; pathogen resistance; plant growth; seed production; gene; ds.

XX Arabidopsis thaliana.

OS

XX WO2003072763-A1.

XX 04-SEP-2003.

XX 28-FEB-2003; 2003WO-CA000274.

XX 28-FEB-2002; 2002CA-02373903.

PR 28-FEB-2002; 2002US-00086464.

XX (GORI/) GORING D.

PA (SILV/) SILVA N.

PA (HAFF/) HAFFANI Y Z.

XX Goring D, Silva N, Haffani YZ;

XX WPI; 2003-712727/67.

XX Producing a transgenic plant having an increased plant resistance, plant growth or seed production comprises transforming a plant with a nucleic acid molecule having a proline-rich Extensin-like Receptor Kinase activity.

XX Disclosure; Fig 21d; 123pp; English.

XX The invention relates to producing a transgenic plant having increased plant height, number of branches, number of seed pods and/or seed production compared to a non-transgenic plant, and/or quicker flowering or later senescence compared to a non-transgenic plant. The method involves transforming a plant with a vector including a proline-rich Extensin-like Receptor Kinase (PERK) nucleic acid molecule or a nucleic acid molecule having PERK activity. The method, as well as the PERK nucleic acid molecule and polypeptide, are useful in increasing plant resistance to wounding and pathogens and in increasing plant growth and seed production. The nucleic acid molecule and polypeptide may also be used in producing transgenic plants or transgenic host cells. The present sequence represents an A. thaliana PERK1 DNA (At3g24550)

XX Sequence 1959 BP; 496 A; 514 C; 425 G; 524 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	5,286-78	Length:	1959
Score:	2936.50	Matches:	574
Percent Similarity:	90.61%	Conservative:	24
Best Local Similarity:	86.97%	Mismatches:	41
Query Match:	85.04%	Indels:	21
DB:	9	Gaps:	12

US-10-086-464-2 (1-647) x ACP36557 (1-1959)

QY 1 MetSerSerAlaProSerProGlyThr-----GlySerProProSerProProSer 17

DB 1 ATGTCCACAGCGCGTCTCCAGGCACCTACTCCATCACCATCTCCACCGCTCTCTCCCA 60

QY 18 AsnSerThrThrThrProProProAla-----SerAlaProProProThrPro 35

DB 61 AACTCGACACACACACCTCTCTCTCCAGCAGCCTCTCTCTCTCTCTCTCTCTCTCTCT 120

QY 36 SerSerProProPro---ProSerThrThrProThrSerProProSerSerArgSer 54

DB 121 TCTTCT 180

QY 55 ThrProSerAlaProProProSerProProProThrProProSerThrProGlySerProPro 74

DB 181 COTCCTTCT 228

QY 75 LeuProGlnProSerProProAlaProThrThrProGlySerProProAlaProValThr 94

DB 229 CTTCTCTCAACCTTCCCTCTCGCTCCCATCACTCTCTCTCTCTCTCTCTCTCTCTCTCT 285

QY 95 ProProThr---ArgAsnProProProSerValProGlyProProProSerAsnProSerArg 113

DB 286 CCCTCAAAACCTCGAAGCCCTCCATCTCTCAACCAAGGACCAACCAACACCTCCCTCA 342

QY 114 GluGlyGlySerProArgProProSerSer-----ProSerProProSerProSerSer 131

DB 343 ---GGATCTACTCTAGAACTCCATCAACACTAAACCGTCGCGCGCTCTGATCTCTTCC 399

QY 132 AspGlyLeuSerThrGlyValValGlyValGlyLeuAlaLeuGlyValAlaLeuVal 151

DB 400 GATGGATTGCTACCGGAGTTGTGTAGGAATCCCATTTGGAGGAGTTCGCTATTCTTGT 459

QY 152 IleValThrLeuIleCysLeuLeuCysLysLysLysArgArgArg-----AspGluGlu 169

DB 460 ATACTGACTCTGATTTGTCTCTCTGTAAAGAAACGAAGAAGACACACACGACGATGAA 519

QY 170 AspAlaTyrTyrValProProProProGlyProGlyProLysAlaGlyGlyProTyrGly 189

DB 520 GCTGCTTACTATGTTCT 579

QY 190 GlyGlnGlnGlnTTPArgGlnGlnAsnAlaThrProProSerAsp---HisValVal 208

DB 580 GGTCAACAACAGTATTGGCAACAACAAACCGTCACGGCGCTCAGATAATCATGTAGTG 639

QY 209 ThrSerLeuProProProProLysAlaProSerProProArgGlnProProProProPro 228

DB 640 ACGTCATTG---CCACCACCTAAGCCTCCATCTCCACCACGAAACCTCTCTCCGCCACT 696

QY 229 ProProPro---PheMetSerSerSerGlyGlySerAspTyrSerAspArgProValLeu 247

DB 697 CCACCACGACATTCATGAGTAGCAGTGGTGGTCTCTGACTATTTCGGATCTTCCGGTCT 756

QY 248 ProProProSerProGlyLeuValLeuGlyPheSerLysSerThrPheThrTyrGluGlu 267

DB 757 CCTCCACCATCTCAGGCGCTTGTGTAGGCTTTCTTAAAGACATTTTCACTATATGAGGAG 816

QY 268 LeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyPheGly 287

DB 817 TTGTCGAGAGCTACTAATGGCTCTCTGAGGCTAATTTGTTAGGACAAAGAGGGTTTGGT 876

QY 288 TyrValHisLysGlyValLeuProSerGlyLysGluValAlaValLysGlnLeuVal 307

DB 877 TATGTGCATAAAGGTATATTGCTAGTGGGAAAGATTGCTGTGAACACACTTGAAGCT 936

QY 308 GlySerGlyGlnGlyGluArgGluPheGlnAlaGluValGluIleIleSerArgValHis 327

DB 937 GGTAGTGTTCAGGAGAGAGAGAGATTTTCAGGCTGAGGTTGAGATCATTAGCAGAGTTCT 996

QY 328 HisArgHisLeuValSerLeuValGlyTyrCysIleAlaGlyAlaLysArgLeuVal 347

DB 997 CACAGGCATTGTTGTTCTCTTATTGTTATTGTTATGCGCGGTGTTCAAAAGATTACTTCT 1056

QY 348 TyrGluPheValProAsnAsnLeuGluLeuHisIleHisGlyGluGlyArgProThr 367

DB 1057 TATGAGTTGTTTCCAAACACAATCTTGAGTTTCCCTCCATGTGTAAGGAGCGGCTACG 1116

QY 368 MetGluTPSerThrArgLeuLysIleAlaLeuGlySerAlaLysGlyLysSerTyrLeu 387

DB 1117 ATGGAATGGAGTACTAGATTGAGATTGCTCTTGAGATCTGCTTAAGGACTTTCATATCTT 1176

QY 388 HisGluAspCysAsnProLysIleIleHisArgAspIleLysAlaSerAsnIleLeu 407

DB 1177 CATGAAGATTGCAATCCGAAATCATTCACCGTGATTAAGGCGTCAACATATTGATT 1236

QY 408 AspPheLysPheGluAlaLysValAlaAspPheGlyLeuAlaLysIleAlaSerAspThr 427

DB 1237 GATTTCAAATTTGAAGCTTGAAGTTGCTGACTTTGCTTCTTGCACCAAGATTGCTTCTG 1296

QY 428 AsnThrHisValSerThrArgValMetGlyThrPheGlyTyrLeuAlaProGluTyrAla 447

```
Db 1297 AACACTCATGTATCTACACGCGTGTATGGAAACCTTTGGGTATTTGGCTCCGGAATAGCT 1356
Qy 448 AlaSerGlyLeuThrGluLeuSerAspValPheSerPheGlyValValLeuLeuGlu 467
Db 1357 GCAAGTGAAGAGCTCACAAAAGTCTGACGTTTCTCATTTGGCGTTGTACTTTTGAA 1416
Qy 468 LeuIleThrGlyArgArgProValAspAlaAsnAsnValTyrValAspSerLeuVal 487
Db 1417 CTTATTACTGGAGGCGCCTGTTGATGCGAACAACTGTCTATGTAGATGACAGCTTAGTT 1476
Qy 488 AspTrpAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGluGlyLeuAla 507
Db 1477 GACTGGGCACCACTGCTTAAACCGAGCATCTCAGGAAGGAGATTTTGGGGTTTGCT 1536
Qy 508 AspAlaLeuMetAsnAsnGlyTyrAspArgGluGluMetAlaArgMetValAlaCysAla 527
Db 1537 GATTCAAGATGGGTATGATGATGACAGAGGAGATGGCTCGCATGTTGCTTGCGCT 1596
Qy 528 AlaAlaCysValArgHisSerAlaArgArgProArgMetSerGlnIleValArgAla 547
Db 1597 GCGGCTTGTTGCCATTGAGCTCGCGCAGACCTCGCATGAGCCAGATAGTACGGCG 1656
Qy 548 LeuGluGlyAsnValSerLeuSerAspLeuAsnGluGlyMetArgProGlyGlnSerAsn 567
Db 1657 TTAGAAGAAATGATCGCTGCTCTCATCTTAACGAAGGGATGAGACCGGGTCACAGCAAC 1716
Qy 568 ValTyrSerSerTyrGlyGlySerThrAspTyrAspSerSerGlnTyrAsnGluAspMet 587
Db 1717 GTATACAGCTCATATGAGGAAGACACACTATGACAGCCCACTCAACGACGACGATG 1776
Qy 588 LysLysPheArgLysMetAlaLeuGlyThrGlnGluTyrAsnAlaThrGlyGluTyrSer 607
Db 1777 ATAAAGTTTAGAATAATGGCTCTTGGAACCTCAAGATACGCCACAAACCGCGAGTACAGT 1836
Qy 608 AsnProThrSerAspTyrGlyLeuTyrProSerGlySerSerSerGluGlyGlnThrThr 627
Db 1837 AATCCAAACCACTAGTACGGAAGTACCGCTGCTGGTTCAAGCAGTGAAGGTCAAGCCACA 1896
Qy 628 ArgGluMetGluMetGlyLysIleLeuArgThrGlyGlnGlyTyrSerGlyProSerLeu 647
Db 1897 CGAGAAATGGAGATGGGAAGATTAAAGAAACCGGTCAAGGTTATAGTGGACCTCTCTT 1956

RESULT 5
ID AAC36968 standard; DNA; 2324 BP.
XX AAC36968;
AC AAC36968;
DT 17-OCT-2000 (first entry)
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 15710.
XX Arabidopsis thaliana.
XX Hybridisation assay; genetic mapping; gene expression control;
XX protein identification; signal transduction pathway; metabolic pathway;
XX promoter; termination sequence; ss.
XX Arabidopsis thaliana.
XX EP1033405-A2.
XX 06-SEP-2000.
XX 25-FEB-2000; 2000EP-00301439.
XX 25-FEB-1999; 99US-0121825P.
XX 05-MAR-1999; 99US-0123180P.
XX 09-MAR-1999; 99US-0123548P.
XX 23-MAR-1999; 99US-0125788P.
XX 25-MAR-1999; 99US-0126264P.
XX 29-MAR-1999; 99US-0126785P.
XX 01-APR-1999; 99US-0127462P.
XX 06-APR-1999; 99US-0128234P.
XX 08-APR-1999; 99US-0128714P.
XX 16-APR-1999; 99US-0129845P.
XX 19-APR-1999; 99US-0130077P.
XX 21-APR-1999; 99US-0130449P.
XX 23-APR-1999; 99US-0130510P.
XX 28-APR-1999; 99US-0130891P.
XX 30-APR-1999; 99US-0131449P.
XX 30-APR-1999; 99US-0132407P.
XX 04-MAY-1999; 99US-0132484P.
XX 05-MAY-1999; 99US-0132485P.
XX 06-MAY-1999; 99US-0132486P.
XX 06-MAY-1999; 99US-0132487P.
XX 07-MAY-1999; 99US-0132863P.
XX 11-MAY-1999; 99US-0134256P.
XX 14-MAY-1999; 99US-0134218P.
XX 14-MAY-1999; 99US-0134219P.
XX 14-MAY-1999; 99US-0134221P.
XX 14-MAY-1999; 99US-0134370P.
XX 18-MAY-1999; 99US-0134768P.
XX 19-MAY-1999; 99US-0134941P.
XX 20-MAY-1999; 99US-0135124P.
XX 21-MAY-1999; 99US-0135353P.
XX 24-MAY-1999; 99US-0135629P.
XX 25-MAY-1999; 99US-0136021P.
XX 27-MAY-1999; 99US-0136392P.
XX 28-MAY-1999; 99US-0136782P.
XX 01-JUN-1999; 99US-0137222P.
XX 03-JUN-1999; 99US-0137528P.
XX 04-JUN-1999; 99US-0137502P.
XX 07-JUN-1999; 99US-0137724P.
XX 08-JUN-1999; 99US-0138094P.
XX 10-JUN-1999; 99US-0138540P.
XX 10-JUN-1999; 99US-0138847P.
XX 14-JUN-1999; 99US-0139119P.
XX 16-JUN-1999; 99US-0139452P.
XX 16-JUN-1999; 99US-0139453P.
XX 17-JUN-1999; 99US-0139492P.
XX 18-JUN-1999; 99US-0139454P.
XX 18-JUN-1999; 99US-0139455P.
XX 18-JUN-1999; 99US-0139456P.
XX 18-JUN-1999; 99US-0139457P.
XX 18-JUN-1999; 99US-0139458P.
XX 18-JUN-1999; 99US-0139459P.
XX 18-JUN-1999; 99US-0139460P.
XX 18-JUN-1999; 99US-0139461P.
XX 18-JUN-1999; 99US-0139462P.
XX 18-JUN-1999; 99US-0139463P.
XX 18-JUN-1999; 99US-0139750P.
XX 18-JUN-1999; 99US-0139763P.
XX 21-JUN-1999; 99US-0139817P.
XX 22-JUN-1999; 99US-0139899P.
XX 23-JUN-1999; 99US-0140353P.
XX 23-JUN-1999; 99US-0140354P.
XX 24-JUN-1999; 99US-0140695P.
XX 28-JUN-1999; 99US-0140823P.
XX 29-JUN-1999; 99US-0140991P.
XX 30-JUN-1999; 99US-0141287P.
XX 01-JUL-1999; 99US-0141842P.
XX 01-JUL-1999; 99US-0142154P.
XX 02-JUL-1999; 99US-0142055P.
XX 06-JUL-1999; 99US-0142390P.
XX 08-JUL-1999; 99US-0142803P.
XX 09-JUL-1999; 99US-0142920P.
XX 12-JUL-1999; 99US-0142977P.
XX 13-JUL-1999; 99US-0143542P.
XX 14-JUL-1999; 99US-0143624P.
XX 15-JUL-1999; 99US-0144005P.
XX 16-JUL-1999; 99US-0144085P.
XX 16-JUL-1999; 99US-0144086P.
XX 19-JUL-1999; 99US-0144325P.
XX 19-JUL-1999; 99US-0144331P.
XX 19-JUL-1999; 99US-0144332P.
```


Qy 170 AspAlaTyrTyrValProProProProProGlyProLysAlaGlyGlyProTyrGly 189
Db 624 GCTGTTACTATGTT---CCTCCTCCTCCTTCTGCTCCCAAGCTGGAGGACCTTACGGT 680
Qy 190 GlyGlnGlnGlnTTPArgGlnGlnAsnAlaThrProProSerAsp---HisValVal 208
Db 681 GGTCAACAACAGTATGTGCAACAACAACCGCTACCGCGCTCAGATAATCATGTAGTG 740
Qy 209 ThrSerLeuProProProProProProProProProProProProProProProPro 228
Db 741 ACGTCAATG---CCACCACTTAAGCTTCATCTCCACCAAGAACTCTCCGCCCACT 797
Qy 229 ProProPro---PheMetSerSerSerGlySerAspTyrSerAspArgProValLeu 247
Db 798 CCACCACCATTCATGAGTAGCAGTGGTTCCTGACTATTCGGATCTTCCGGTCTT 857
Qy 248 ProProProSerProGlyLeuValLeuGlyPheSerLysSerThrPheThrTyrGluGlu 267
Db 858 CCTCCACCATCTCCAGGCTTGTGTAGGCTTTTCAAAAAGCACTTTCACTTATAGGAG 917
Qy 268 LeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyPheGly 287
Db 918 TTGTCAAGACTACTAATGGCTTCTCTAGGCTAATTTGTAGCAAGGAGGTTTGGT 977
Qy 288 TyrValHisLysGlyValLeuProSerGlyLysGluValAlaValLysGlnLeuLysVal 307
Db 978 TATGTGCATAAGGTATATTCCTAGTGGGAAAGAGTTGCTGTGAAACAGTTGAAAGCT 1037
Qy 308 GlySerGlyGlnGlyGluArgGluPheGlnAlaGluValGluLysLeuSerArgValHis 327
Db 1038 GGGAGTGGTCAGGAGAGAGAGAGTTTCAGGCTGAGGTTGAGATCATTAGCAGAGTTTCA 1097
Qy 328 HisArgHisLeuValSerLeuValGlyTyrCysIleAlaGlyAlaLysArgLeuVal 347
Db 1098 CACAGGATTTGGTTTCTTATTGGTTATGTATGGCTGATGTCAAGATTCCTTGTGTC 1157
Qy 348 TyrGluPheValProAsnAsnLeuGluLeuHisLeuHisGlyGlyArgProThr 367
Db 1158 TATGAGTTTGTCTCAACAACATCTTGAGTTTCACTCCATGTTAAGGGAAGGCTTACA 1217
Qy 368 MetGluTrpSerThrArgLeuLysIleAlaLeuGlySerAlaLysGlyLeuSerTyrLeu 387
Db 1218 ATGGAATGGAGTACTAGATTGAAGATTGCTCTTGATCTGCTAAAGGACTTTTCATATCTT 1277
Qy 388 HisGluAspCysAsnProLysIleIleHisArgAspIleLysAlaSerAsnIleLeuIle 407
Db 1278 CATGAGATTGCAATCCGAAATCATTCACCGTATATTAAGGCTCAACATATTGATT 1337
Qy 408 AspPheLysPheGluAlaLysValAlaAspPheGlyLeuAlaLysIleAlaSerAspThr 427
Db 1338 GATTTCAAAATTTGAAGCTAAGTTGCTGACTTTGCTTGGCAGATTGCTTCTGTATACA 1397
Qy 428 AsnThrHisValSerThrArgValMetGlyThrPheGlyTyrLeuAlaProGluTyrAla 447
Db 1398 AACACTCATGTATCTACACGCGTGTGAGAACCTTTGGGTATTTGGCTCCAGATATGCT 1457
Qy 448 AlaSerGlyLysLeuThrGluLysSerAspValPheSerPheGlyValValLeuLeuGlu 467
Db 1458 GCAAGCGGAAGCTCACAGAAAAGTCTGACGTTTCTCATTTGGTGTGTACTTTTGGAG 1517
Qy 468 LeuIleThrGlyArgArgProValAspAlaAsnAsnValTyrValAspSerLeuVal 487
Db 1518 CTTATACCGGGAGCGCCCTTGTGATGCAACAATGCTATGTATGATGACAGCTTAGTT 1577
Qy 488 AspTrpAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGlyLeuAla 507
Db 1578 GACTGGGCACGACCTTGTCTTAACCGGATCTCTGAGGAAGGAGATTTTGTGTTGGCT 1637
Qy 508 AspAlaLysMetAsnAsnGlyTyrAspArgGluLeuMetAlaArgMetValAlaCysAla 527
Db 1638 GATTCAAGAAGTGGTAATAGTAGTAG-AGAGAGAGAGATGCTCCATGTGGTGTGGCT 1696
Qy 528 AlaAlaCysValArgHisSerAlaArgArgProArgMetSerGlnIleValAlaAla 547

Db 1697 GCAGCTTGTGTGCTCCATTTCAGCTCGCGCAGACCTCGCATGAGCCAGATAGTACGGCG 1756
Qy 548 LeuGluGlyAsnValSerLeuSerAspLeuAsnGluGlyMetArgProGlyGlnSerAsn 567
Db 1757 TTAGAAGGAATGTATCGCTGTCTGATCTTAAACGAAGGGATGAGACCGGGTCAAGCAAC 1816
Qy 568 ValTyrSerSerTyrGlySerThrAspTyrAspSerSerGlnTyrAsnGluAspMet 587
Db 1817 GTATACAGCTCATATGGAGGAGCCTGACTATGACACAGGCAATCAACAGATGACATG 1876
Qy 588 LysLysPheArgLysMetAlaLeuGlyThrGlnGluTyrAsnAlaThrGlyGlyTyrSer 607
Db 1877 AAGAAGTTTAGGAAATGGCTCTTGGAACTCAAGATACGCAACCCGGGAGTACAGT 1936
Qy 608 AsnProThrSerAspTyrGlyLeuTyrProSerGlySerSerGlyGlnThrThr 627
Db 1937 AATCCAAACCACTGACTACCGACTGTACCCGCTCTGGTTCAAGCAGTGAAGGTCAAGCCACA 1996
Qy 628 ArgGluMetGluMetGlyLysIleLysArgThrGlyGlnGlyTyrSerGlyProSerLeu 647
Db 1997 CGAAGATGGAGATGGGAAGATTAGAAAACCGTCAAGGTTATAGTGGACCTCTCTT 2056
RESULT 6
ACF36554
ID ACF36554 standard; DNA; 1812 BP.
XX AC ACF36554;
DT 18-DEC-2003 (first entry)
XX DE O. sativa PERK protein encoding genomic DNA.
XX KW PERK1; transgenic; plant; proline-rich extensin-like receptor kinase;
XX KW wound; pathogen resistance; plant growth; seed production; rice; gene;
XX KW ds.
OS Oryza sativa.
XX WO2003072763-A1.
XX PD 04-SEP-2003.
XX PF 28-FEB-2003; 2003WO-CA000274.
XX PR 28-FEB-2002; 2002CA-02373903.
XX PR 28-FEB-2002; 2002US-00086464.
XX (GORI/) GORING D.
PA (SILV/) SILVA N.
PA (HAFF/) HAFFANI Y Z.
XX Goring D, Silva N, Haffani YZ;
XX WPI; 2003-712727/67.
XX
XX Producing a transgenic plant having an increased plant resistance, plant
XX growth or seed production comprises transforming a plant with a nucleic
XX acid molecule having a Proline-rich Extensin-like Receptor Kinase
XX activity.
XX
XX Disclosure; Fig 21a; 123pp; English.
XX
XX The invention relates to producing a transgenic plant having increased
XX plant height, number of branches, number of seed pods and/or seed
XX production compared to a non-transgenic plant, and/or quicker flowering
XX or later senescence compared to a non-transgenic plant. The method
XX involves transforming a plant with a vector including a proline-rich
XX Extensin-like Receptor Kinase (PERK) nucleic acid molecule or a nucleic
XX acid molecule having PERK activity. The method, as well as the PERK
XX nucleic acid molecule and polypeptide, are useful in increasing plant
XX resistance to wounding and pathogens and in increasing plant growth and
XX seed production. The nucleic acid molecule and polypeptide may also be

Db 1140 ATGGAATGAGAAAGATGAAG---AACAAATCAAGGTTTTCAGTGAAGTTCT 1187
 RESULT 9
 AAF77096
 ID AAF77096 standard; DNA; 2520 BP.
 XX
 AC
 XX AAF77096;
 DT 17-MAY-2001 (first entry)
 XX
 DE Arabidopsis gene #2.
 XX
 KW Proline-rich extensin-like receptor kinase; PERK; resistance; plant; ds.
 XX
 OS Arabidopsis thaliana.
 XX
 FN WO200114563-A1.
 XX
 PD 01-MAR-2001.
 XX
 PR 19-AUG-1999; 99US-0149466P.
 PR 13-OCT-1999; 99US-0159122P.
 XX
 PA (GORI/) GORING D.
 PA (SILV/) SILVA N.
 XX
 FI Goring D, Silva N;
 XX
 DR WPI; 2001-244305/25.
 XX
 PT New proline-rich, extensin-like receptor kinase nucleic acids and
 PT polypeptides useful for increasing plant wounding or pathogen resistance,
 PT or for producing transgenic plants with increased wounding or pathogen
 PT resistance.
 XX
 PS Example; Fig 12; 91pp; English.
 XX
 CC The present invention relates to proline-rich extensin-like receptor
 CC kinase (PERK). The PERK nucleic acids and polypeptides are useful for
 CC increasing the resistance of plants to wounding and pathogens. These are
 CC also useful for producing transgenic plants with increased wounding and
 CC pathogen resistance compared with a wild type plant, as well as in assays
 CC for identifying and developing compounds to inhibit and/or enhance
 CC polypeptide function directly
 XX
 SQ Sequence 2520 BP; 717 A; 553 C; 537 G; 713 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 2,93e-36 Length: 2520
 Score: 1507.00 Matches: 368
 Percent Similarity: 53.47% Conservative: 79
 Best Local Similarity: 44.02% Mismatches: 146
 Query Match: 43.64% Indels: 244
 DB: 4 Gaps: 22
 US-10-086-464-2 (1-647) x AAF77096 (1-2520)
 Qy 3 SerAlaProSerProGlyThrGlySerProProSerProProSerAsnSerThrThr 22
 Db 36 TCTGCTCTCCACAAACCTCCACCTCTCTCCATCTCCACCGCTCAATACCAATTCACCC 95
 Qy 23 ThrProProAlaSerAlaProProProThrThrProSerSerProProProSer 42
 Db 96 ACCTCTTCTCGG-----CCGGCTCCGCTCTCTCTCTCTCTCTCTCTCTCTCTCAAGGA 146
 Qy 43 ThrileProThrSerProProSerSerArgSerThrProSerAlaProProProSer 62
 Db 147 GACTCATCATCATCGCCACCTCTCTGATTCACATCTCCACCGCTCCACAGCTCTCTAAC 206
 Qy 63 ProProThrProSerThrProGlySerProProProLeuProLeuProSerProAla 82

Db 207 CTTCTTAAT---TCCTCTAATAACTCTCTCTCCCTCCGTCACAGGGCGGTGGAGGAA 263
 Qy 83 ProThrThrProGlySerProProAlaProValThrProProThrArgAsnProProPro 102
 Db 264 AGAGGAAATGGAGAAACAATGGTGCAATGATCTCCCGTCCAGGGCTCTCTCTCT 323
 Qy 103 SerValProGlyProProSerAsnProSerArgGlyGlySerProArg-----Pro 120
 Db 324 TCT-----CCTCTCTTAGGAGTAATGGAGATAATGGTGTAGCAGATCATCGCCA 374
 Qy 121 ProSer-----SerProSerProProSerPro-----SerSerAsp 132
 Db 375 CCAGGAGACACTGGAGGCTCTCGCTCAGACAACCTCTCTCTAGCGGAGGAGCAGTGA 434
 Qy 133 Gly-----LeuSerThrGlyValValGlyValGlyIleAlailegly 145
 Db 435 GGAGGTGGAGGTGAAGAAGTAATACGATACAGCGATCATAGTTGGTATTAGTCGGA 494
 Qy 146 GlyValAlaLeuLeuValIleValThrLeuLeuCysLeuLeuCysLysLysLysArgArg 165
 Db 495 GCTGGACTTTTGTATGATCGTTCTTATTATTGTGTCTTAGACGCCAANAAGAGAGAAA 554
 Qy 166 ArgAspGluGluAspAlaTyrTyrValProProProProProProGlyProLysAla--- 184
 Db 555 -----GACTCTCTCTAC-----CCTGAACCCATGAAGGTAAACATATAC 596
 Qy 185 -----GlyGlyPro 187
 Db 597 ACACCTTTATGTTTCAACAAATAAGAGCTTAGATTCTTTCAATAAATTTTCAGAAACCA 656
 Qy 188 TyrGlyGlyGlnGlnGlnTrpArgGlnGlnAsnAlaThrProProSerAspHisVal 207
 Db 657 ATATCAATACTATGGAACCAACACACAACTATGTTTACACAGATTTATCCGAA----- 710
 Qy 208 ValThrSerLeuProProProProLysAlaProSerProPro----- 221
 Db 711 -----TTGGCACCTAAATTCACAGGCCAAACCAACAACTACTGTGTGTGGG 761
 Qy 222 ArgGln-ProProProProProProPheMetSerSerSergly-----GlySe 239
 Db 762 AGCGGTGGACCATCATCGGCTCTCTCCGCGATGCTTACAGCGGAGAGATTCTTTC 821
 Qy 239 rAspTyrSer-----AspArgProValLeuProProProSerProGlyLeuValLeuGl 257
 Db 822 CATGTACTCAGGCCCATCATCGCCAGTTTACCTCTCTCTCGCTCTCTAGCCCTCGG 881
 Qy 257 yPheSerLysSerThrPheThrTyrGluGluLeuAlaArgAlaThrAsnGlyPheSerGl 277
 Db 882 ATTCAACAGAGCAGCTTTTACTTACCAAGAGCTTGGCGCTGCAACAGGAGGTTTACGGA 941
 Qy 277 uAlaAsnLeuLeuGlyGlnGlyGlyPheGlyTyrValHisLysGlyValLeuProSerGl 297
 Db 942 TGCTTAACCTTTTGGGACAGGAGATTGGGTATGTCATAAAGAGCTTGTCCCTAGCGG 1001
 Qy 297 YLysGluValAlaValLysGlnLeuLysValGlySerGlyGlnGlyGluArgGluPheGl 317
 Db 1002 GAAAGAAGTAGCAGTTAAGAGTTTAAAGCGGGTAGCGGACAGGAGAGGAGTTTCA 1061
 Qy 317 nAlaGluValGluIleLeuSerArgValHisHisArgHisLeuValSerLeuValGlyTy 337
 Db 1062 AGCTGAGGTGCATATCATTAGCGGTGTCATCATCGGTATCTTGTCTTCTTGTGGATA 1121
 Qy 337 rCysIleAlaGlyAlaLysArgLeuLeuValTyrGluPheValProAsnAsnLeuGl 357
 Db 1122 TTGCATAGCTGATGGACAGAGGATGTGGTTTATGAGTTTGTTCCTTAAACAACTTTGGA 1181
 Qy 357 uLeuHisLeuHisgly----- 362
 Db 1182 ATATCATCTTCATGGTTAGACCACTTAAACACTTTGAGTACTAAGTTTATTTCTTAAT 1241
 Qy 363 -----Glu-GlyArgProThrM 368
 Db 1242 CTATATATTCAAGAAAGTTGTAACCTTAAATTTGTTGTTAGGAGAAATAATCTTCCGATA 1301

```
QY 368 etGluTrpSerThrArgLeuIysIleAlaLeuGlySerAlaIysGlyLeuSerTyrLeuH 388
Db 1302 TGGAGTTCTCCACTAGTTGGGTATCGGCTTAGGTGCTCGGAAAGGACTCGTTACCTTC 1361
QY 388 isGluAspCys----- 391
Db 1362 ACGAAGACTGTAAGTTTAAACATTCCACCTTCATCTCTATTTTCTTAACCAAGTTGCATAAAAC 1421
QY 392 -----Asn 392
Db 1422 AGAGAAAGCTCTGCTCTGACTAGTGTATCTTTTTGGCTGAGAAATGTTGCAGGCCAT 1481
QY 393 ProLysIleIleHisArgAspIleLysAlaSerAsenIleLeuIleAspPheLysPheGlu 412
Db 1482 CCTCGGATCATTACCGGACATCAAGTCTGCAATATTTCTTTGGGACTTCAACTTTGAT 1541
QY 413 Ala----- 413
Db 1542 GCTATGGTGATAAACTAGTAGCTTCATCTCATCTGCTTTTGTAAAGACTACATTGA 1601
QY 414 -----LysValAlaAspPheGlyLeuAlaLysIleAlaSerAs 426
Db 1602 TGACATTTTGCAATTTGTTTATTCAGGTGCTGATTTTGGATTAGTAACTAATCACTGA 1661
QY 426 pThrAsnThrHisValSerThrArgValMetGlyThrPhe-Gly----- 440
Db 1662 TAAACAACACTCATGTATCTACTCGTGTGATGGAACTTTTCGGGTAAAGCGTTTACCGTAT 1721
QY 440 ----- 440
Db 1722 GATAAGATTGTCGTGACACTCAAGAAACATAACCTTTGTAGACTAATCTACTTTGCTTC 1781
QY 441 -----TyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrG 454
Db 1782 TTTCCACAAACATGTGTAGATATCTAGCTCCAGATATGCTTCAAGCGGTAAATTAACCG 1841
QY 454 LuLysSerAspValPheSerPheGlyValValLeuLeuLeuIleThrGlyArgArgP 474
Db 1842 AGAAATCCGATGTTTTCTCTTACGGAGTTATGTTTATGGAATTAATACTGGAAACGAC 1901
QY 474 roValAspAlaAsnValTyrValAspAspSerLeuValAspTrp----- 489
Db 1902 CGGTTGAT---AATAGCATCACCATGGACGACACCTTAGTAGATTGGGTATTCATGATG 1958
QY 489 ----- 489
Db 1959 TAACATATGATCGGTATATATGTTTTTCGGCTTTTCGGCTACTAATGATCATGAATA 2018
QY 490 ---AlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGluGlyLeuAlaAsp 508
Db 2019 CAGGCTCGGCTCTTTATGCTCGCGCTAGAGATGGAACCTTTAATGAGCTCGCAGAT 2078
QY 509 AlaLysMetAsnAsnGlyTyrAspArgGluGluMetAlaArgMetValAlaCysAlaAla 528
Db 2079 GCGAGGCTTGAAGGCAACTACAAACCCGCAAGAAATGGCTCGAATGGTACTTTGTCGCCCT 2138
QY 529 AlaCysValArgHisSerAlaArgArgProArgMetSer----- 542
Db 2139 GCTAGCATTCGTCATTCGGGGCTTAACGCTCCAAAGATGAG-CCAGGTGAATCAAAATTA 2197
QY 543 -----GlnIleVal 545
Db 2198 TAACTAAAGTCTATTTTTTGTGAGAGAATAACAAACAATGTTGGTTTTTCAGATAGTA 2257
QY 546 ArgAlaLeuGluGlyAsnValSerLeuSerAspLeuAsnGluGlyMetArgProGlyGln 565
Db 2258 AGAGCGTTAGAGAGGAGAGTGTCTTAGATGCTTTTAACGAAGGTGTGAAGCCAGGACAC 2317
QY 566 SerAsnValTyrSerSerTyrGlyGlySerThrAspTyrAspSerSerGlnTyrAsnGlu 585
Db 2318 ACTAAGCTTTACGGGTCTATTGGGAGCAAGCTCGGATTATAGTCAGACATCTTTACAATGCA 2377
```

```
QY 586 AspMetLysLysPheArgLysMetAlaLeuGlyThrGlnGluTyr-----AsnAla 602
Db 2378 GACATGAAGAAATTCAGACAGATAGCTTTTGTGAGCAAGAATTTCCAGTCAGTACTGT 2437
QY 603 ThrGlyGluTyrSerAsnProThrSerAspTyrGly 614
Db 2438 GAAGGACATCTAGTATGATTCTTAGAGATATGGGA 2473

RESULT 10
ACF36551
ID ACF36551 standard; DNA; 2520 BP.
XX AC
AC ACF36551;
XX 18-DEC-2003 (first entry)
DT Arabidopsis PERK1 receptor related protein encoding DNA.
DE PERK1; transgenic; plant; proline-rich extensin-like receptor kinase;
KW wound; pathogen resistance; plant growth; seed production; gene; ds.
XX Arabidopsis thaliana.
OS WO2003072763-A1.
PN 04-SEP-2003.
PD 28-FEB-2003; 2003WO-CA000274.
XX 28-FEB-2002; 2002CA-02373903.
PR 28-FEB-2002; 2002US-00086464.
XX (GORI/) GORING D.
PA (SILV/) SILVA N.
PA (HAFF/) HAFFANI Y Z.
XX Goring D, Silva N, Haffani YZ;
XX WPI; 2003-712727/67.
DR P-PSDB; ABR82940.
XX
PT Producing a transgenic plant having an increased plant resistance, plant
PT growth or seed production comprises transforming a plant with a nucleic
PT acid molecule having a Proline-rich Extensin-like Receptor Kinase
PT activity.
PS Disclosure; Fig 12; 123pp; English.
XX
CC The invention relates to producing a transgenic plant having increased
CC plant height, number of branches, number of seed pods and/or seed
CC production compared to a non-transgenic plant, and/or quicker flowering
CC or later senescence compared to a non-transgenic plant. The method
CC involves transforming a plant with a vector including a proline-rich
CC Extensin-like Receptor Kinase (PERK) nucleic acid molecule or a nucleic
CC acid molecule having PERK activity. The method, as well as the PERK
CC nucleic acid molecule and polypeptide, are useful in increasing plant
CC resistance to wounding and pathogens and in increasing plant growth and
CC seed production. The nucleic acid molecule and polypeptide may also be
CC used in producing transgenic plants or transgenic host cells. The present
CC sequence represents a PERK1 polypeptide related protein encoding genomic
CC DNA from A. thaliana (Accession NO. AAD15491)
XX
SQ Sequence 2520 BP; 717 A; 552 C; 538 G; 713 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 3.58e-36 Length: 2520
Score: 1504.00 Matches: 367
Percent Similarity: 53.47% Conservative: 80
Best Local Similarity: 43.90% Mismatches: 146
Query Match: 43.56% Indels: 224
DB: 9 Gaps: 22

US-10-086-464-2 (1-647) x ACF36551 (1-2520)
```


Db 987 CCTGTGCTGGTCTT-----CACAGTGATGCTTCCAACTTAACCGGT 1028
Qy 244 ArgProValLeuProProSerProGlyLeuValLeuGlyPheSerLysSerThrPhe 263
Db 1029 CGAACTGCTATACCGTCACTCAA---GCTGCAACTCTTGGTCACAAACCAAGACACTTTC 1085
Qy 264 ThrTyrGluGluLeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGln 283
Db 1086 ACATACGATGAATGCTCCATTGCAACAGAGGTTTCGCTCAAGTCAAATTTGCTAGGACAA 1145
Qy 284 GlyGlyPheGlyTyrValHisGlyGlyValLeuProSerGlyLysGluValAlaValLys 303
Db 1146 GGAGGATTTGGGTATGTTTCAATAAGGAGTCTGCTAGTGGCAAGAGTTGCGAGTGAAG 1205
Qy 304 GlnLeuLysValGlySerGlyGlnGlyValArgGluPheGlnAlaGluValGluLeuLeu 323
Db 1206 AGCTTTAAACTTGAAGTGGACAGGGAAGCGAGTTTCAAGCAGAGTTGATATCATTT 1265
Qy 324 SerArgValHisHisArgHisLeuValSerLeuValGlyTyrCysIleAlaGlyAlaLys 343
Db 1266 AGCGGTGTCCATCATCGTTCATCTGTTCTCTGTTGGATATTGCATCTCTGGTGGTCAA 1325
Qy 344 ArgLeuLeuValTyrGluPheValProAsnAsnLeuLeuGluLeuHisLeuHis----- 361
Db 1326 AGACTTTTGGTTTATGAGTTTATACCTTAACAACACTCTTGAATTTTCATCTCATGTACA 1385
Qy 361 ----- 361
Db 1386 TTCATCTAACAGAAATGTTTCTTGTTATTAACAAACCTTTAAAGTATGTTTCTCTTTAAT 1445
Qy 362 -----GlyGluGlyArgProThrMetGluTrpSerThrArgLe 374
Db 1446 CAGGAACATGATTGAAATTTTCAGGAAGGTCGTCCGGTTTGGATTGGCCTACAGAGT 1505
Qy 374 uLysIleAlaLeuGlySerAlaLysGlyLeuSerTyrLeuHisGluLeuAspCys----- 391
Db 1506 GAAGATTGCATTGGGATCAGCTAGAGGCTTGCATATTGTCATGAAGACTGTGAAGAAAT 1565
Qy 391 ----- 391
Db 1566 CTTTATCTCACATATTTGTCATCAGTTCTATCTCGCTCTTACAAATATTGAAAGATTGT 1625
Qy 392 -----AsnProLysIleIleHisArgAspIleLysAlaSerAs 404
Db 1626 ATATGTTACATCAATTATAGTCACTCCCTCGCATTTCCACAGATATCAAGCTGCAAA 1685
Qy 404 nIleLeuLeuAspPheLysPheGluAla----- 413
Db 1686 CATCTCTTCTGATTTTCAGTTTGTAGAC-CAAGGTATGTGTGTATATATATCGACTCTTGTAC 1744
Qy 413 ----- 413
Db 1745 TACTTTTACTTTCATGTCCTCATATTTTGTTCCTCAATCTGTCGATGTGTATCAGT 1804
Qy 414 -----LysValAlaAspPheGlyLeuAlaLysIleAlaSerAspTh 427
Db 1805 CTTATTGTGTAATAATATGACAGTGGCAGATTTTGGATTGGCTAAGCTATCTCAAGACAA 1864
Qy 427 rAsnThrHisValSerThrArgValMetGlyThrPhe-Gly----- 440
Db 1865 CTATACTCATGTCCTCACTCGCGTCATGGGAACCTTTTGGGTAAAGCAGCTTTGTAAAAATGT 1924
Qy 440 ----- 440
Db 1925 CTCAACTCATCCACACTTATTTAGTTTCTTTCATCTGTTTAAACATTTTCTTGGATTCA 1984
Qy 441 --TyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPheS 460
Db 1985 GATACTTAGTCCAGAGTATGATCAACGGGAAGATTATCCGACAAATCTGATGTTTCT 2044
Qy 460 erPheGlyValValLeuGluLeuIleThrGlyArgArgProValAspAlaAsnAsnV 480
Db 2045 CATTTGGAGTAATGCTTCTTGGAGTTCATTAACCGGAAGACCTCTCTGATCTAACTCGA- 2103

Qy 480 alTyrValAspAspSerLeuValAspTrp----- 489
Db 2104 --GAAATGGGAAGATAGCTTGGTAGATTGGGTAAAGTCCCGCCTCTTCGGTTTACTTT 2161
Qy 489 ----- 489
Db 2162 GTTTAATCCCAAAACACTTTCCAAAGCAAAACAGAAACAAATCTTACTATTGTGTGTC 2221
Qy 490 --AlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGluGlyLeuAlaAspA 509
Db 2222 AGCAAGGCTTTGTGTTTGAAGCAGCTCAAGATGGAGATTACAAACAATTTGCTGATC 2281
Qy 509 lAlysMetAsnAsnGlyTyrAspArgGluGluMetAlaArgMetValAlaCysAlaAlaA 529
Db 2282 CACGCTCAGAGCTAAACTCAGTCATCAAGAGATGGTTCAAATGGCTTCTTGTGACAGTG 2341
Qy 529 laCysValArgHisSerAlaArgArgProArgMetSerGlnIle----- 544
Db 2342 CAGCAATCAGACATTCAGCAAGAAGACGGCTTAAGATGAGCCAGGT-TCAAAAAACTCATA 2400
Qy 544 ----- 544
Db 2401 CCACCTTGTGTTCTATTGTTATATTTTACTCACAATTAATCTTGATGATAAATGTGAC 2460
Qy 545 -----ValArgAlaLeuGluG 550
Db 2461 ATACTAATGAATCTTGAACATGTGTATGTTAAATGAAGATTGTACGAGCACTAGAG 2520
Qy 550 lyAsnValSerLeuSerAspLeuAsnGluGlyMetArgProGlyGlnSerAsnValTyrS 570
Db 2521 GAGATATGTCATATGATGATCTAAGTGAGGGAACAACACAGGACAAAGCAGCTACTTGA 2580
Qy 570 erSerTyrGlyGlySerThrAspTyrAspSerSerGlnTyrAsnGluAspMetLysLysP 590
Db 2581 GCCCGGAGCGTCAGCTCAGAGTATGACGCAAGCTCGTACACGGCAGACATGAAAAAGT 2640
Qy 590 heArgLysMetAlaLeuGlyThrGlnGluTyrAsnAlaThrGlyGlyTyrSerAsnProT 610
Db 2641 TCAAGAAACCTGGGCTTAGAGAAATAAGAAATATCAAGCAGT---GAATATGGTGA---A 2694
Qy 610 hrSerAspTyrGlyLeuTyrProSerGlySerSerGlyGlnThrThrArgGluM 630
Db 2695 CAAGTGATGATGCTTAAACCTTCTGCTTCAAGTAGTGAA-----GAAA 2739
Qy 630 etGluMetGlyLysIleLysArgThrGlyGln 640
Db 2740 TGAATAGAGGCTCAATGAACGCAATCTCTCAG 2771
RESULT 12
ACF36552
ID ACF36552 standard; DNA; 2820 BP.
XX AC ACF36552;
XX AC
DT 18-DEC-2003 (first entry)
XX
DE Arabidopsis PERK1 receptor related protein encoding DNA.
XX PERK1; transgenic; plant; proline-rich extensin-like receptor kinase;
KW wound; pathogen resistance; plant growth; seed production; gene; ds.
XX Arabidopsis thaliana.
XX
XX WO2003072763-A1.
XX
XX 04-SEP-2003.
XX
XX 28-FEB-2003; 2003WO-CA000274.
XX
XX 28-FEB-2002; 2002CA-02373903.
PR 28-FEB-2002; 2002US-00086464.
XX

PA (GORI/) GORING D.
 PA (SILV/) SILVA N.
 PA (HAFF/) HAFFANI Y Z.
 XX
 PI Goring D, Silva N, Haffani YZ;
 DR WPI; 2003-712727/67.
 DR P-PSDB; ABR82941.
 XX
 XX Producing a transgenic plant having an increased plant resistance, plant
 PT growth or seed production comprises transforming a plant with a nucleic
 PT acid molecule having a Proline-rich Extensin-like Receptor Kinase
 PT activity.
 XX
 PS Disclosure; Fig 13; 123pp; English.
 XX
 PS The invention relates to producing a transgenic plant having increased
 CC plant height, number of branches, number of seed pods and/or seed
 CC production compared to a non-transgenic plant, and/or quicker flowering
 CC or later senescence compared to a non-transgenic plant. The method
 CC involves transforming a plant with a vector including a Proline-rich
 CC Extensin-like Receptor Kinase (PERK) nucleic acid molecule or a nucleic
 CC acid molecule having PERK activity. The method, as well as the PERK
 CC nucleic acid molecule and polypeptide, are useful in increasing plant
 CC resistance to wounding and pathogens and in increasing plant growth and
 CC seed production. The nucleic acid molecule and polypeptide may also be
 CC used in producing transgenic plants or transgenic host cells. The present
 CC sequence represents a PERK1 polypeptide related protein encoding genomic
 CC DNA from A. thaliana (Accession NO. CMA18823)
 XX
 SQ Sequence 2820 BP; 819 A; 630 C; 576 G; 795 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 3,42e-35 Length: 2820
 Score: 1471.50 Matches: 366
 Percent Similarity: 49.89% Conservative: 99
 Best Local Similarity: 39.27% Mismatches: 152
 Query Match: 42.62% Indels: 318
 DB: 9 Gaps: 28

US-10-086-464-2 (1-647) x ACP36552 (1-2820)

Qy 1 MetSerAlaProSerProGlyThrClySerProProSerProSerAsnSerThr 20
 Db 52 GTGGATTTCATCTCTGCTCCCTGGAACCC---TCAATGGGACACCCACCTGCAACGGAACA 108
 Qy 21 ThrThrProProAlaSerAlaProProThrProThrProSerSerProPro 40
 Db 109 TCG-----CCGTCTAATGAGTCATCGCGCCACACCACTCTCTTCCACCAACCA 159
 Qy 41 ProSerThrIleProThrSerProProSerSerArgSerThrProSerAlaProPro 60
 Db 160 TCATCA-----ATATCTGCTCCTCGGCAGATATCTCCGCTCTCTTTTTCACCGCG 210
 Qy 61 ProSerProProThr-----ProSerThrProGlySerProProProLeu 75
 Db 211 CCTGACCACCAACGCAAGAAACGTCACCTCTACATCTCTGCTCCTCATCGCGCGCTGT 270
 Qy 76 -----ProGln-----ProSerProAlaPro-----Thr 84
 Db 271 GTAGTAATCCGTCAACCGCAGACTCCAGAGAAATCTTCTCCACCTGCACTGAAGGCTCA 330
 Qy 85 ThrProGlySerProProAla---ProValThrPro-----ProThrArgAsn 99
 Db 331 ACTCTGTAACGCGCACTGCAACCAACCAACACCGTCGAACCAATCACCGGAAAGACCA 390
 Qy 100 ProProPro-----Pro----- 102
 Db 391 ACTCTCTCTTCTCTCTGGTGCCTAATGATGACCGAAACAGAACCAATGGCGGAAACCAAC 450
 Qy 103 -----SerValProGlyProProSerAsnProSerArg-----GluGly 116
 Db 451 AGAGACGCTCCACACCATCACCCCGTCGTCAGGGAACAGAACTTCCGCTGACGGTGC 510

Qy 117 SerProArgProProSerSerProSerProSerProSerSerSerSerSerSerSerThr 136
 Db 511 TCACCTTTCACACCTCGGTGATTAAGCCCTCCCTCAGAAATAGTGAGGATTCAGACTCATCA 570
 Qy 137 -----GlyValValValGlyIleAlaIleGlyGlyVal 147
 Db 571 TCGGTAATCATCACCAAGCCACATTTGGATTGATTGAGTCCCTTGTAGGACGAGG 630
 Qy 148 AlaLeuLeuValIleValThrLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 167
 Db 631 CTTTGTCTTCTACTTGCAGTGTGTATTTGCATCTGTTCGCAACAGCAAGAGAGAGAGAAA 690
 Qy 168 GluGluAspAla-----TyrTyrValProProProProProProProProGlyProLeuAla 184
 Db 691 TCTCCTCAGGTCACCAACCATGCACTACTACAAATAACAATCCTTATGGAGGAGACCCCTCA 750
 Qy 185 Gly----- 185
 Db 751 GGTAAATTACAGTTTAGTATTAACCTGGAATTTAATTTGTAGCTAATGGTGTGTTGATTAGGT 810
 Qy 186 -----GlyProTyrGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 194
 Db 811 TTCAGAACGATCATAGTCTAATGCTTCTGCTAGCTCCATATGCGCAAAAGGATTAGATT- 869
 Qy 195 TrpArgGlnGlnAsnAlaThr-----ProPro 203
 Db 870 TATAAGCTAAAGAGAGATGTTGTCATAGTGTAGTAAATGGTGGTATTATTAAGGGGAACACCT 929
 Qy 204 SerAspHisValValThrSerLeuProProProProProProProProProProProPro 223
 Db 930 CAAGATCATGTGGTG---AATATGGCTGGTCAAGGAGTGGGAATGGGGTCCACAGCAA 986
 Qy 224 ProProProProProProProProPheMetSerSerSerSerSerSerSerSerSerSerSer 243
 Db 987 CCTGTGTCTGCTCT---CACAGTGAATGCTTCCCACTTAACCGGT 1028
 Qy 244 ArgProValLeuProProProProSerProGlyLeuValLeuGlyPheSerLysSerThrPhe 263
 Db 1029 CGAACTGCTATACCGTCACCTCAA---GCTGCAACTCTTGGTCAACCAAGACACTTTC 1085
 Qy 264 ThrTyrGluGluLeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGln 283
 Db 1086 ACATACATGACATGCTCCATTCGACAGAGAGTTTCCTCAGTCAATTTGCTAGGACAA 1145
 Qy 284 GlyGlyPheGlyTyrValHisLysGlyValLeuProSerGlyLysGlyValAlaValLys 303
 Db 1146 GGAGGATTTGGGTATGTTTCAAAAGAGATTCTGCTAGTGGCAAGAAAGTTGCAGTGAAG 1205
 Qy 304 GlnLeuLysValGlySerGlyGlnGlyGluArgGluPheGlnAlaGluValGluIle 323
 Db 1206 AGTCTTAACTTGAAGTGGACCAAGGGGAACGCGAGTTTCAAGCAGAGTTGATATCAT 1265
 Qy 324 SerArgValHisHisArgHisLeuValSerLeuValGlyTyrCysIleAlaGlyAlaLys 343
 Db 1266 AGCGTGTCCATCATCTGCTTCTCTTGTGGATATTCGCACTCTCTCGTGGTCAA 1325
 Qy 344 ArgLeuLeuValTyrGluPheValProAsnAsnAsnLeuGluLeuHisLeuHis----- 361
 Db 1326 AGACTTTGGTTTATGAGTTTATACCTTAACCAACTCTTGAATTTTCATCTTCATGTTACA 1385
 Qy 361 ----- 361
 Db 1386 TTCACTTAACAGAAATGTTTCTTGTATTAACAAACCTTTAAGTATGTTTCTCTTTAAT 1445
 Qy 362 -----GlyGluGlyArgProThrMetGluTrpSerThrArgLe 374
 Db 1446 CAGGAACATGATTGAATTTTCAGGAAGGGTCTCGTCCGTTTGGATTGGCTCACAAGAGT 1505
 Qy 374 uLysIleAlaLeuGlySerAlaLysGlyLeuSerTyrLeuHisGluAspCys----- 391
 Db 1506 GAAGATTGCAATTGGGATCAGCTAGAGGCTTGCATATTTGCATGAAGAACTGTAAGAAAT 1565


```
Qy 391 ----- 391
Db 1566 CTTTATCTCACATATTGTCATCAGTTTCTATCTCGCTCTCTACAAATATTTGAAAGATTGT 1625
Qy 392 -----
Db 1626 ATATGTTACATCAATTATAGTCACCTCGCATTTATCCACAGAGATATCAAGCTGCAAA 1685
Qy 404 nileuileaspPheLysPheGluAla----- 413
Db 1686 CATCTCTTGATTTTCAGTTTGTAGAC-CAAGGTATGTGTATATATCGACTCTTGTAC 1744
Qy 413 ----- 413
Db 1745 TACTTTTACTTTTCATTTGCTCTCTATTTTGTTCCTCAATCTGTGTCGATGTGTATCAGT 1804
Qy 414 -----LysValAlaAspPheGlyLeuAlaLysLeuAlaLysLeuAlaSerAspTh 427
Db 1805 CTTATTGTGTAATATATGTCAGGTGGCAGATTTTGGATTGGCTAAGCTATCTCAAGACAA 1864
Qy 427 rAsnThrHisValSerThrArgValMetGlyThrPhe-Gly----- 440
Db 1865 CTATACTCATGTCCTCCACTCCGCTCATCGGAACTTTTGGGTAAAGCAGCTTTGTAATGT 1924
Qy 440 ----- 440
Db 1925 CTCAACTCATCCACACTTATTAGTTTCTTCTTCACTTGTGTTTAAACATTTTCTTGGATTCA 1984
Qy 441 --TyrLeuAlaProGluTyrAlaLaserGlyLysLeuThrGluLysSerAspValPheS 460
Db 1985 GATACTTAGCTCCAGAGATGTCATCAAGCGGAAAGTTATCCGACAAATCTGATGTTTCT 2044
Qy 460 erPheGlyValValLeuLeuLeuLeuThrGlyArgGlyProValAspAlaAsnAsnV 480
Db 2045 CATTTGGAGTAATGCTTCTGAGCTCATACCGGAAGACCTCTCTGGAATCACTGGA- 2103
Qy 480 alTyrValAspSerLeuValAspTrp----- 489
Db 2104 -GAAATGGAAGATAGCTTGTAGATTGGTAAAGTTCGGTCCCGCTCTTCGGTTTACTT 2161
Qy 489 ----- 489
Db 2162 GTTTAATCCCAAAACACTTTCCAAAGCAAAACAGAAACAAATCTTACTATTGTTGTC 2221
Qy 490 --AlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGlyLeuAlaAspA 509
Db 2222 AGCAAGCCCTTTGTTGTTGAAGAGCGCTCAAGATGGAGATTAACCAATTCGCTGATC 2281
Qy 509 laLysMetAsnAsnGlyTyrAspArgGluGluMetAlaArgMetValAlaCysAlaAlaA 529
Db 2282 CACGCTAGACCTAAACTACAGTCATCAAGAGATGGTTCAATGGCTCTTGTGCAGCTG 2341
Qy 529 laCysValArgHisSerAlaArgArgProArgMetSerGlnIle----- 544
Db 2342 CAGCAATCAGACATTCAGCAAGAAGACGGCTTAAGATGAGCAGCGGT-TCAAAAACTCATA 2400
Qy 544 ----- 544
Db 2401 CCACCTGTTGTTCTTATGTTATATTTTACTCAATTAATCTTGATGATAAATGTGAC 2460
Qy 545 -----ValArgAlaLeuGluG 550
Db 2461 ATACTAATGAATCTTGAAACATGTGTATGGTAAATGAAAGATTGTACGAGCACTAGAG 2520
Qy 550 lyAsnValSerLeuSerAspLeuAsnGluGlyMetArgProGlyGlnSerAsnValTyrS 570
Db 2521 GAGATATGTCATGATGATCTAAGTGAAGGAAACAAGACCAAGCAGGACGACTTCTGA 2580
Qy 570 erSerTyrGlySerThrAspTyrAspSerSerGlnTyrAsnGluAspMetLysLysP 590
Db 2581 GCCCGGAGCGGTGAGCTCAGAGTATGACCAAGCTCGTACCGCGGACATGAAAGT 2640
Qy 590 heArgLysMetAlaLeuGlyThrGlnGluTyrAsnAlaThrGlyGluTyrSerAsnProT 610
```

```
Db 2641 TCAGAAACTGGCGTTAGAGATAAAGATATCAAGCAGT---GAATATGTTGGA---A 2694
Qy 610 hrSerAspTyrGlyLeuTyrProSerGlySerSerGlyGlnThrThrArgGluM 630
Db 2695 CAAGTGAATAGCTTAAACCCCTCTCTCTCAAGTAGTGAA-----GAAA 2739
Qy 630 etGluMetGlyLysIleLysArgThrGlyGln 640
Db 2740 TGAATAGAGGCTCAATGAACGCAATCCTCAG 2771
RESULT 13
AAF77095
ID AAF77095 standard; DNA; 2880 BP.
XX
AC AAF77095;
DT 17-MAY-2001 (first entry)
XX
DE Arabidopsis gene #1.
XX
KW Proline-rich extensin-like receptor kinase; PERK; resistance; plant; ds.
XX
OS Arabidopsis thaliana.
XX
FN W0200114563-Al.
PD 01-MAR-2001.
XX
PF 18-AUG-2000; 2000WO-CA000966.
XX
PR 19-AUG-1999; 99US-0149466P.
PR 13-OCT-1999; 99US-0159122P.
XX
(GORI/) GORING D.
PA (SILV/) SILVA N.
XX
Goring D, Silva N;
XX
WPI; 2001-244305/25.
XX
New proline-rich, extensin-like receptor kinase nucleic acids and
polypeptides useful for increasing plant wounding or pathogen resistance,
or for producing transgenic plants with increased wounding or pathogen
resistance.
XX
Example; Fig 11; 91pp; English.
XX
The present invention relates to proline-rich extensin-like receptor
kinase (PERK). The PERK nucleic acids and polypeptides are useful for
increasing the resistance of plants to wounding and pathogens. These are
also useful for producing transgenic plants with increased wounding and
pathogen resistance compared with a wild type plant, as well as in assays
for identifying and developing compounds to inhibit and/or enhance
polypeptide function directly
XX
SQ Sequence 2880 BP; 814 A; 693 C; 569 G; 804 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 1.5e-27 Length: 2880
Score: 1209.50 Matches: 325
Percent Similarity: 43.85% Conservative: 92
Best Local Similarity: 34.17% Mismatches: 160
Query Match: 35.03% Indels: 376
DB: 4 Gaps: 26
US-10-086-464-2 (1-647) x AAF77095 (1-2880)
Qy 2 SerSerAlaProSerProGlyThrGlySerProProSerProProSerAsnSerThrThr 21
Db 98 AACTCAGCTCTTCCACCTGTTGATTCCTCTCTCTAGTCCACGCTGATTCATCA 157
Qy 22 Thr-----ThrProProAlaSer----- 28
```

Db 158 ACACGCCGCTGTCAGAACCATCCACTCTCTCCAGATTCCAGCTTCCTCCTTACCT 217
|||
Qy 29 -----AlaProProThrThrProSerSerProProPro 40
:::|
Db 218 TCGATTCTTCCTCCGCTAACAGATTCTCCACTCCACCTCCGATTCTTCCTCCACCGTT 277
:::|
Qy 41 ProSerThrIleProThrSerProProSerSerArgSerThrProSerAlaPro--- 59
|||
Db 278 GATTCAACC---CCTTCTCCGCGCCACCGAGCTCAACGAATCTCTCTCTCTCCACGAA 334
|||
Qy 60 -----ProProSerProProThrProSerProGlySerProProPro--- 74
|||
Db 335 GATTCCGAAACACACCACTCTCTCCACCAATGAATCCAAATGACCAACACCTCTCTCCGCT 394
|||
Qy 75 -----LeuProGlnProSerProProLa-----ProThrThrPro 86
|||
Db 395 CAAGATCTTCAATCGCTCTCCATCGTCGCGCTCGCGAATGTAGGACCCACAAACCG 454
|||
Qy 87 GlySerProPro-----AlaProValThrProProThrArgAsnPro-----Pro 101
|||
Db 455 GAATCACCACCGTTACAACTCTCTCCAGCTCCACCGATCAGATCTTACAAATTCACCG 514
|||
Qy 102 ProSerValProGlyProProSerAsnPro-----SerArgGluGlyGlySerPro 118
|||
Db 515 CCAGCTTACCATTAGACCTACCAATCTCTCCCAATACAAACCATCAGGACCGCACT 574
|||
Qy 119 ArgProProSerProPro-----SerProProSerPro----- 129
|||
Db 575 TCTCTCCGGCTAATCCCAAGCTCCGCGCGGCCATTCCCAACAGTACCACCCCAAACT 634
|||
Qy 129 ----- 129
|||
Db 635 CTTCTAGTGGACCTGTGGTGTCTCCATCTCTCACATCCCTAGTAAAGAACTCCTACT 694
|||
Qy 130 -----SerSerAspGlyLeuSerThrGly-----ValVal 139
|||
Db 695 CCMAACCAAGCAATGGAGATGGCGGTGGCGGTGTGGCGGTATCAAGGAAGACTATG 754
|||
Qy 140 ValGlyIleAlaIleGlyGlyValAlaLeuLeuValIleValThrLeuIleCysLeuLeu 159
|||
Db 755 GTTGTATGGCTGTAGCGGTTTCGCAATCATGGCGCTTATAGGCGTGTGTCTTAGTG 814
|||
Qy 160 CysIleLysLysArgArgAspGluGluAspAlaTyr-----TyrValPro 175
|||
Db 815 AGAAGAAAGAAAGAGA-----AACATTGATGCTATAATCACTCACAGTACTTGCCA 868
|||
Qy 176 ProPro----- 177
|||
Db 869 CATCCCAATTTCTGTGTAATCATAGGTTTAAATCTCAACCTTATCTCTCTGTGATCAT 928
|||
Qy 177 ----- 177
|||
Db 929 CTTCTATGTGTTGAATCATCTCTGACTATCTTTGCTTTTGTATGTAGTGGATTCTTA 988
|||
Qy 178 -----ProProProGlyProLysAlaGlyProTyrGlyGlyGlnGlnGln 194
|||
Db 989 TACGGTCAGATCCAGGTAAGGATPACTCCTCTGCTCTGCTTAATGGTTCAATGATAACAAT 1048
|||
Qy 195 TrpArgGlnGlnAsnAlaThrProProSerAspHisValThrSerLeuProProPro 214
|||
Db 1049 TCACAGCAACAACATCTCTATGGGAACAGTTATGTGTACAGCTGGT----- 1096
|||
Qy 215 ProLysAlaProSerProProArgGlnProProProProProProPheMetSer 234
|||
Db 1097 -----GGTGGTTATCTCTCATCATCA-----ATGCAA 1123
|||
Qy 235 SerSerGlyGlySerAspTyrSerAspArgProValLeuProProProProSerProGlyLeu 254
|||
Db 1124 TCAAGTGGACACCTGC-----TCTGCT 1147
|||
Qy 255 ValLeuGlyPheSerLysSerThrPheThrTyrGluGluLeuAlaArgAlaThrAsnGly 274
|||

Db 1148 ATACTCGGAAGTGGCCAGACTCAATTCAGTTACGAAGAGCTTGCTGAGATAACAACAGGC 1207
|||
Qy 275 PheSerGluAlaAsnLeuLeuGlyGlnGlyGlyPheGlyTyrValHisLysGlyValLeu 294
|||
Db 1208 TTTCTCGCAAAAACATTTCTGGAGAAAGCGGATTTGGATGTGTCTATAAAGGTACATTG 1267
|||
Qy 295 ProSerGlyLysGluValAlaValLysGlnLeuLysValGlySerGlyGlnGlyGluArg 314
|||
Db 1268 CAGATGGTAAAGTTGTTGCGTTAAGCAGCTTAAAGCTGGAAGTGGACAGGTGACCGT 1327
|||
Qy 315 GluPheGlnAlaGluValGluIleSerArgValHisHisArgHisLeuValSerLeu 334
|||
Db 1328 GAATTCAGACAGAGTTGAGATCATCAGCGCGTTTCATCATCGCATTTGCTCTCTG 1387
|||
Qy 335 ValGlyTyrCysIleAlaGlyAlaLysArgLeuLeuValTyrGluPheValProAsnAsn 354
|||
Db 1388 GTTGGTTACTGCAATTCAGACCAGCATAGATTGCTTATCATATGATGATGTTTCTAATCAA 1447
|||
Qy 355 AsnLeuGluLeuHisLeuHisGlyGlu----- 363
|||
Db 1448 ACCTGGAGCATCATTTGCGTGGTGGTACTGTTGTTACCATTTTCGTTATAGATAAGACT 1507
|||
Qy 364 -----GlyArgProTh 367
|||
Db 1508 TTTTGTAGCTTTACGTGTTAGACTGACTCGCTTTACGCTTTAGGAAAGGTTTGCAGT 1567
|||
Qy 367 rMetGluTyrSerThrArgLeuLysIleAlaLeuGlySerAlaLysGlyLeuSerTyrLe 387
|||
Db 1568 TTTAGATGGTCTAAGAGAGTCCGGATCGGTATAGGATCAGCCAAAGGTTGGCATATCT 1627
|||
Qy 387 uHisGluAspCysAsn----- 392
|||
Db 1628 TCAGAAAGCTGTAA--GTAATGCCTTCACATTTTCTTAGTTGTGTGCTTTGGTTATGCAC 1686
|||
Qy 393 -----ProLy 394
|||
Db 1687 TTTATAGTTTAAACAGAAAGCCAAAATCATATCTCTGTTTATTTTACAGTCTATCCGAA 1746
|||
Qy 394 sIleIleHisArgAspIleLysAlaSerAsnIleLeuIleAspPheLysPheGluAlaLy 414
|||
Db 1747 AATCATTCACAGATATATAAGTCAAGCAATATTTCTTCTAGATGATGAATATGAAGCTCA 1806
|||
Qy 414 s----- 414
|||
Db 1807 -GGCAATAATGAATCTCTCTTTTCGTTAAATCTATCTTATGACTGTAAAGTTTAGTTA 1865
|||
Qy 415 -----ValAlaAspPheGlyLeuAlaLysIleAl 424
|||
Db 1866 ATGAGACTTGTCTGTTTTTTGGATGTTTAGTTGCTGATTTTGGACTTGTAGACTCAA 1925
|||
Qy 424 aSerAspThrAsnThrHisValSerThrArgValMetGlyThrPhe-Gly----- 440
|||
Db 1926 TGATAACAACAACAACTCATGTTCACCTCGGGTTATGGGAACCTTCGGGTAAAGCAACAT 1985
|||
Qy 440 ----- 440
|||
Db 1986 TCATCACAACCTTACTCCAAAACCTGGACCTTATTGATCCAATGCCCTGATGAAAAGTTTG 2045
|||
Qy 440 ----- 440
|||
Db 2046 TTATATATGGCTTGAGGCAACAAATTTGGATCAAACTGAAATCTTTTATTTGATCGTATGGCT 2105
|||
Qy 441 -----TyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuT 453
|||
Db 2106 GCATGACATGTTTGTGTTAAGGTACCTAGCGCGGAATATGTCATCAAGTCGGAATTTGA 2165
|||
Qy 453 hrGluLysSerAspValPheSerPheGlyValValLeuLeuGluLeuIleThrGlyArg 473
|||
Db 2166 CTGATAGATCGATGTATTCTCATTCGGGGTTGTTCTCTTAGAGCTTGTAACTGACCGGA 2225
|||
Qy 473 rgProValAspAlaAsnAsnValTyrValAspAspSerLeuValAspTrp----- 489
|||
Db 2226 AACCAGTTGACAGACTCAGCCTCTAGGAGAAGAGAGTTTGGTTGTAATGGGTAAAGAAATCC 2285
|||

QY 489 ----- 489
 Db 2286 AACTTTCAACATCTTCAATAATAGTAGATTGGCCCTAGTATATCTATATATAGTACTTA 2345
 QY 490 -----AlaArgProLeuLeuAenArgAlaSerGluGlnGlyAspPheGlu 504
 Db 2346 TAAATGAACACACAGCGCGCGCTGCTTCTCAAGCCATTGAGACCGGAGATTTAAGC 2405
 QY 505 GlyLeuAlaAspAlaLysMetAsnAsnGlyTyrAspArgGluGluMetAlaArgMetVal 524
 Db 2406 GAACGTATTGATACACGCGCTTGAAGCGTTATGTGGAGCATGAAGTCTTCAGAAATGATC 2465
 QY 525 AlaCysAlaAlaCysValArgHisSerAlaArgArgArgProArgMetSerGln--- 543
 Db 2466 GAGACAGCGCGTGCATGTGTAGACATCTTGTGTCCTCAAAACGCCACGATGGTTGAGTA 2525
 QY 543 ----- 543
 Db 2526 ATTCTGACTAACCAAAAGTCCAAAGTCCCATATATAGTAACAAGTGATTTCTCACATCT 2585
 QY 544 -----IleValArgAlaLeuGluGlyAsnValSerLeu 554
 Db 2586 GAAACTTATCTACTCTTCGAAATAAGGTTGTGAGACATGGACTCGCGGAGACTCG 2645
 QY 555 SerAspLeuAenGluGlyMetArgProGlyGlnSerAenValTyrSerSerTyrGlyGly 574
 Db 2646 GGAGATATTACACACGGNATCAAAATTTGGCAA----- 2678
 QY 575 SerThrAspTyrAspSerSerGlnTyrAsnGluAspMetLysPheArgLysMetAla 594
 Db 2679 AGCACAACTTATGACTCAGGGCAATACAATGAAGACATTATGAAATTCAGGAAATGGCG 2738
 QY 595 LeuGlyThrGlnGluTyrAsnAlaThrGlyGluTyrSerAenProThrSerAspTyrGly 614
 Db 2739 TTGGTGGTGATACACGGTGAGTCAGGATGTACAGT-----GGA 2780
 QY 615 LeuTyrProSerGlySerSerSerGlu 623
 Db 2781 AACTACTCTGCCAAAGCTCTTCAGAT 2807
 RESULT 14
 ID ACF36550
 AC ACF36550 standard; DNA; 2880 BP.
 AC ACF36550;
 XX 18-DEC-2003 (first entry)
 DE Arabidopsis PERK1 receptor related protein encoding DNA.
 KW PERK1; transgenic; plant; proline-rich extensin-like receptor kinase;
 KW wound; pathogen resistance; plant growth; seed production; gene; ds.
 XX Arabidopsis thaliana.
 XX W02003072763-A1.
 XX 04-SEP-2003.
 XX 28-FEB-2003; 2003WO-CA000274.
 XX 28-FEB-2002; 2002CA-02373903.
 XX 28-FEB-2002; 2002US-00086464.
 XX (GORI/) GORING D.
 XX (SILV/) SILVA N.
 XX (HAFF/) HAFFANI Y Z.
 XX Goring D, Silva N, Haffani YZ;
 XX WPI; 2003-712727/67.
 DR P-PSDB; ABR82939.

XX Producing a transgenic plant having an increased plant resistance, plant
 PT growth or seed production comprises transforming a plant with a nucleic
 PT acid molecule having a Proline-rich Extensin-like Receptor Kinase
 PT activity.
 XX Disclosure; Fig 11; 123pp; English.
 XX The invention relates to producing a transgenic plant having increased
 CC plant height, number of branches, number of seed pods and/or seed
 CC production compared to a non-transgenic plant, and/or quicker flowering
 CC or later senescence compared to a non-transgenic plant. The method
 CC involves transforming a plant with a vector including a Proline-rich
 CC Extensin-like Receptor Kinase (PERK) nucleic acid molecule or a nucleic
 CC acid molecule having PERK activity. The method, as well as the PERK
 CC nucleic acid molecule and polypeptide, are useful in increasing plant
 CC resistance to wounding and pathogens and in increasing plant growth and
 CC seed production. The nucleic acid molecule and polypeptide may also be
 CC used in producing transgenic plants or transgenic host cells. The present
 CC sequence represents a PERK1 polypeptide related protein encoding genomic
 CC DNA from A. thaliana (Accession NO. AAC98010)
 XX
 SQ Sequence 2880 BP; 814 A; 693 C; 569 G; 804 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 1.5e-27 Length: 2880
 Score: 1209.50 Matches: 325
 Percent Similarity: 43.85% Conservative: 92
 Best Local Similarity: 34.17% Mismatches: 160
 Query Match: 35.03% Indels: 376
 DB: 9 Gaps: 26

US-10-086-464-2 (1-647) x ACF36550 (1-2880)

QY 2 SerSerAlaProSerProGlyThrGlySerProProSerProProSerAsnSerThrThr 21
 Db 98 AACTCAGCTCTTCACCTGTTGATTCTCTCTCTAGTCCACGAGTGTATCATCA 157
 QY 22 Thr-----ThrProProAlaSer----- 28
 Db 158 ACACGGCGGTGTGAGAACCATCCACTCTCTCCAGATTCAGATTCCTCTTTTACCT 217
 QY 29 -----AlaProProThrThrProSerSerProPro 40
 Db 218 TCGATTCTCTCCGCTAACAGATTCTCCACTCCACTTCGATTCTCTCCACCGTT 277
 QY 41 ProSerThrIleProThrSerProProSerSerArgSerThrProSerAlaPro--- 59
 Db 278 GATTCAACC---CCTTCTCGCGCCACCGCGTCAAAACGAATCTCTCTCTCCAGAA 334
 QY 60 -----ProProSerProThrProSerThrProGlySerProPro--- 74
 Db 335 GATTCCGAAACACCCACTGCTCTCCACCAATGAATCCAAATGACAAACCCCTCTCCGCT 394
 QY 75 -----LeuProGlnProSerProAla-----ProThrThrPro 86
 Db 395 CAAGATCTTCAATCGCTCTCTCCATCGTCGCGTGGCGAATGTAGGACCCACAAACCG 454
 QY 87 GlySerProPro-----AlaProValThrProThrArgAenPro-----Pro 101
 Db 455 GAATCACCACCGTTTACAATCTCTCCAGTCCACCGATCAGATCATCAAAATTCACCG 514
 QY 102 ProSerValProGlyProProSerAenPro-----SerArgGluGlyGlySerPro 118
 Db 515 CCAGCTTCACCATTAGACCTTACCAATCTCTCCAGTCCACCGATCAGATCATCAAAATTCACCG 574
 QY 119 ArgProProSerSerPro---SerProSerPro----- 129
 Db 575 TCTCTCCGGCTAATCCCAACGCTCCCGCGGAGCCATTCCCCACAGTACCACCCAAACT 634
 QY 129 ----- 129
 Db 635 CTTTCTAGTGGACCTGTGGTGTCTCCATCTCTCACATCCCTCTAGTAAAGGAATCTCTACT 694

```
QY 130 -----SerSerAspGlyLeuSerThrGly-----ValVal 139
Db 695 CCAAAACCAAGCAATGGAGATGGCGGTGGCGGTATCAAGGGAAGACTATG 754
QY 140 ValGlyIleAlaIleGlyValAlaLeuLeuValIleValThrLeuIleCysLeuLeu 159
Db 755 GTTGGTATGGCTGTAGCGGTTTCGAATCATAGCGCTTATAGCGGTGTGTCTTAGTG 814
QY 160 CysIlySlyValArgArgAspGluGluAspAlaTyr-----TyrValPro 175
Db 815 AGAAGAGAGAAAAGAGA-----AACATTGATAGCTATATCACTCACAGTACTTGCCA 868
QY 176 ProPro----- 177
Db 869 CATCCCAATTCTCTGTATAATCAGGTTTAAAAATCTCACCTTTATCTCTCTCATCAT 928
QY 177 ----- 177
Db 929 CTTCTATGTGCTTGAATCATCTCTGACTATCTTTTGTGTTGTAGATGGATTTCTTA 988
QY 178 -----ProProProGlyProlysAlaGlyGlyProTyrGlyGlyGlnGlnGln 194
Db 989 TACGGTCAAGATCCAGGTAAAGGATCTCTCTGTCTCTAATGTTCAATGTATAACAAT 1048
QY 195 TrpArgGlnGlnAsnAlaThrProProSerAspHisValValThrSerLeuProPro 214
Db 1049 TCACAGCAACAACAATCTCTATGGAACAGTTATGTGTACAGCTGGT----- 1096
QY 215 ProlysAlaProSerProProArgGlnProProProValLeuProProProPheMetSer 234
Db 1097 -----GGTGGTATCCTCATCATCAA-----ATGCCAA 1123
QY 235 SerSerGlyGlySerAspTyrSerAspArgProValLeuProProProSerProGlyLeu 254
Db 1124 TCNAGTGCACACCTGAC-----TCTGCT 1147
QY 255 ValLeuGlyPheSerLysSerThrPheThrTyrGluGluLeuAlaArgAlaThrAsnGly 274
Db 1148 ATACTCGGAAGTGGCCAGCTCATTTACGTTACGAAGAGCTTGCTGAGATAACAACAGGC 1207
QY 275 PheSerGluAlaAsnLeuGlyGlnGlyPheGlyTyrValHisLysGlyValLeu 294
Db 1208 TTTGCTCCGAAAAACATTTCTGGAGAGCGGATTTGGATGTGTTCTATAAAGGTACATTTG 1267
QY 295 ProSerGlyLysGluValAlaValLysGlnLeuLysValGlySerGlyGlnGlyGluArg 314
Db 1268 CAGGATGGTAAAGTTGTCGGTTAAGCAGCTTAAAGCTGGAAGTGGACNAGGTGACCGT 1327
QY 315 GluPheGlnAlaGluValGluIleIleSerArgValHisHisArgHisLeuValSerLeu 334
Db 1328 GAAATCAAAGCAGAGGTTGAGATCATCAGCGCGTTTCATCATCGCATTTGCTCTCTG 1387
QY 335 ValGlyTyrCysIleAlaGlyAlaLysArgLeuLeuValTyrGluPheValProAsnAsn 354
Db 1388 GTTGGTTACTGCATTTACAGACAGCATAGATTGCTTATCTATGATGATGTTTCTTAATCAA 1447
QY 355 AsnLeuGluLeuHisHisGlyGlu----- 363
Db 1448 ACCTTGGAGCATCATTTGCATGGTGAGTGACTTTGTACCATTTTCGTTATAGATAAGACT 1507
QY 364 -----GlyArgProTh 367
Db 1508 TTTTATTTAGCTTTACGTGTAGACTGACTCGCTTTTACGCTTTAGGAAAGGTTTGGCCAGT 1567
QY 367 rMetGlnTrpSerThrArgLeuLysIleAlaLeuGlySerAlaLysGlyLeuSerTyrIle 387
Db 1568 TTTAGAGTGGTCTAAGACAGTCCCGATCGCTATAGGATCAGCAAGGGTTGGCATATCT 1627
QY 387 uHisGluAspCysAsn----- 392
Db 1628 TCACGAGACATGTA -GTAATGCCCTTCACATTTTCTTAGTTGTGTGCTTTGTTATGCAC 1686
```

```
QY 393 -----ProLy 394
Db 1687 TTCATAGTTTAAACAGAGCCAAATAATCATATCCTTGTGTTTATTTTACAGGTCATCCGAA 1746
QY 394 sIleIleHisArgAspIleLysAlaSerAsnIleLeuIleAspPheLysPheGluAlaLy 414
Db 1747 ANTCAATTCACAGATATTAAGTCAGCAATATTTCTTCTAGATGATGAATATGAAGCTCA 1806
QY 414 s----- 414
Db 1807 -GGCAATAATGAAATCCTCTTTTCGTTAAATCTATCTTATGACTGTAAAGTTTAGTTA 1865
QY 415 -----ValAlaAspPheGlyLeuAlaLysIleAl 424
Db 1866 ATGAGACTTGTCTGTTTTTTTGGATGTTAGTTGCTGATTTTGGACTTGTAGACTCAA 1925
QY 424 aserAspThrAsnThrHisValSerThrArgValMetGlyThrPhe-Gly 440
Db 1926 TGATACACACAACTCATGTTTCACTCGGGTTATGGAACTTCGGGTAAAGCAACAT 1985
QY 440 ----- 440
Db 1986 TCATCACAACACTCTACTCCAAACTCGACCTTATTGATCCAATGCTGATGAAAGATTG 2045
QY 440 ----- 440
Db 2046 TTATATATGGCTTGAGCAACAAATTGGATCAAACTGAAATCTTTATTTGATCGTATGGCT 2105
QY 441 -----TyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuT 453
Db 2106 GCATGACATGTTTGTGTTAAGTTACCTAGCGCCGGAATATGCATCAAGTGGAATTTGA 2165
QY 453 hrGluLysSerAspValPheSerPheGlyValValLeuLeuGluLeuIleThrGlyArgA 473
Db 2166 CTGATAGATCGATGTTCTCTCAATTCGGGGTTGTTCTCTTAGAGCTTGTAACTGGACGGA 2225
QY 473 rgProValAspAlaAsnAsnValTyrValAspAspSerLeuValAspTrp----- 489
Db 2226 AACCAGTTGACACAGACTCAGCCTCTAGGAGAAGAGAGTTTGGTTGAATGGGTAAAGATCC 2285
QY 489 ----- 489
Db 2286 AACTTTCAACATCTTCAATAATAGTAGATTGGCCCTAGTATATATATATAGTACTTA 2345
QY 490 -----AlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGlu 504
Db 2346 TAAATGAACCTCACAGGCGCGCGCTGCTTCTCAAAGCCATTGAGACCGGAGATTTAAGC 2405
QY 505 GlyLeuAlaAspAlaLysMetAsnAsnGlyTyrAspArgGluGluMetAlaArgMetVal 524
Db 2406 GAACTGATTGATACACGCGTTTGAAGCGTTTATGTGAGCATGAAGTCTTCAGAAATGATC 2465
QY 525 AlaCysAlaAlaLysValArgHisSerAlaArgArgProArgMetSerGln--- 543
Db 2466 GAGACGCGCTGCATGTGTGTAGACATCTCGTCCAAACGCCACGATGTTTCAGGTA 2525
QY 543 ----- 543
Db 2526 ATTCTGACTAACCAAAAGTCCAAAGCTCCCATATATAGTAACAAGTGATTTCTCACATCT 2585
QY 544 -----IleValArgAlaLeuGluGlyAsnValSerLeu 554
Db 2586 GAAAACTTATCTACTCTTCGAAATAAGGTTGTGAGACATTTGGACTGCGCGGAGACTCG 2645
QY 555 SerAspLeuAsnGluGlyMetArgProGlyGlnSerAsnValTyrSerSerTyrGlyGly 574
Db 2646 GGAGATATACACCGGAAATCAAAATTGGGCAA----- 2678
QY 575 SerThrAspTyrAspSerSerGlnTyrAsnGluAspMetLysLysPheArgLysMetAla 594
Db 2679 AGCACACTTATGACTCAGGGCAATACAATGAAGACATTTATGAAATTCAGGAAATGGCG 2738
QY 595 LeuGlyThrGlnGluTyrAsnAlaThrGlyLeuTyrSerAsnProThrSerAspTyrGly 614
```

Db 2739 TTTGGTGGTGATAACAGCGTAGAGTCAGGATTGACAGT-----GGA 2780

Qy 615 LeuTyrProSerGlySerSerSerGlu 623

Db 2781 AACTACTCTGCCAAAGCTCTTCAGAT 2807

RESULT 15

RESOLUT
ACF36553

ACF36553
ID ACF36553 standard; DNA; 3060 BP.

AC ACP36553:

DT 18-DEC-2003 (first entry)

DE Arabidopsis PERK1 receptor related protein encoding DNA.

PERK1: transgenic: plant: proline-rich extensin-like receptor kinase.

KW wound: pathogen resistance: plant growth: seed production: gene: ds: FERK1; crabsyemc; plant; protein-11c1 extensin-like receptor kinase

OS *Arabidopsis thaliana*.

AA PN WO2003072763-A1.

04-SEP-2003.

XX
PF 28-FEB-2003: 2003WO-CA0000274.XX
PR 28-FEB-2002: 2002CA-02373903-

PR 28-FEB-2002; 2002US-00086464.
XX

PA (GORI//) GORING D.
PA (STIV//) STIVA N

PA (SILV/) SILVA N.
PA (HAFF/) HAFFANT

XX
XX
XXXX / / (mm)
XXXXXX

PI Goring D, Silva N, Haffani YZ;

DR WPI; 2003-712727/67.

DR P-PSDB; ABR82942.

PT Producing a transgenic plant having an increased plant resistance, plant growth or seed production comprises transforming a plant with a nucleic acid molecule having a Proline-rich Extensin-like Receptor Kinase PT activity.

PS Disclosure: Fig 14: 123pp: English.

The invention relates to producing a transgenic plant having increased plant height, number of branches, number of seed pods and/or seed production compared to a non-transgenic plant, and/or quicker flowering or later senescence compared to a non-transgenic plant. The method involves transforming a plant with a vector including a proline-rich Extensin-like Receptor Kinase (PERK) nucleic acid molecule or a nucleic acid molecule having PERK activity. The method, as well as the PERK nucleic acid molecule and polypeptide, are useful in increasing plant resistance to wounding and pathogens and in increasing plant growth and seed production. The nucleic acid molecule and polypeptide may also be used in producing transgenic plants or transgenic host cells. The present sequence represents a PERK1 polypeptide related protein encoding genomic DNA from *A. thaliana* (Accession NO. CAA118590)

Sequence 3060 BP: 797 A: 766 C: 592 G: 905 T: 0 U: 0 Other:

Alignment Scores:

Pred. No.:	8.1e-27	Length:	3060
Score:	1185.00	Matches:	342
Percent Similarity:	39.28%	Conservative:	72
Best Local Similarity:	32.45%	Mismatches:	185
Query Match:	34.32%	Indels:	456
DB:	9	Gaps:	31

US-10-086-464-2 (1-647) x ACF36553 (1-3060)

[illegible]

[illegible]

2122	Db	GTTTAAAAATATTAATTTACGGTTTCAAGTTTCAAGTTTCAAGTTTCAAGTTTCTTCCTT	2181
440	Qy	-----	440
2182	Db	GTTGTATCGTAATCCAGATCAAAAGATTATTTGATTACTAAATGCTTGTGTGCACCGTCT	2241
441	Qy	-----TyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuTh	453
2242	Db	ATTTGGTATGCATTTAAACACAGATACATGGCTCCAGAATACGGCTCCAGTGGTAAAGTAAC	2301
453	Qy	rGluLysSerAspValPheSerPheGlyValValLeuLeuGluLeuIleThrGlyArgAr	473
2302	Db	TGATAAATCAGATGATATATTCCTTTGGGTCTGTCCTTAGAACTCATCTCTGAGCGTCC	2361
473	Qy	gProValAspAlaAsnAsnValTyrValAspAspSerLeuValAspTrp	489
2362	Db	ATCAATTTTCGCCCAAGATTCTTCCACAACACAGAGTTTAGTAGACTGGGTAAGTCAAAG	2421
489	Qy	-----	489
2422	Db	TACATGATGATGATGATACCATTAGGTTTCTTTTCCCACTAGTTATTATAGATGAAA	2481
489	Qy	-----	489
2482	Db	ATGAACTCCAGTATTAATCTCAAGCTTGAAACTTCATTAGACTGTAAATTTTGATTATTC	2541
490	Qy	-----AlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGluGlyLeuAla	508
2542	Db	CTCAGGCGAGGCCATTGCTTTACGAAGCAATCTCTGAAGAAGTCTTTTGACTTTCTTTGTAG	2601
508	Qy	spAlaLysMetAsnAsnGlyTyrAspArgGluGluMetAlaArgMetValAlaCysAla	528
2602	Db	ACTCAAGTTTGGAGAAGAATTACGATCAACTCAGATGGCAACATGGCTGTGTGCTG	2661
528	Qy	laLaCysValArgHisSerAlaArgArgProArgMetSerGlnIle	544
2662	Db	CTGCTGCATCGCCAACTAGCTTGGCTTCGGCTAGATGAGCCAGGT-CTGAGATTTA	2720
544	Qy	-----	544
2721	Db	GTTTAAATACATGTATTTTCGTCCATATATCCGAAGGAAACAAAGCTAATCCATCAATACAT	2780
545	Qy	-----ValArgAlaLeuGluGlyAsnValSerLeus	555
2781	Db	TTATATCTGAAAAAACTTGAATAGTAGTACGTGCTCTTGAAGGCGAGGTGGCCCTGA	2840
555	Qy	erAspLeuAsnGluGlyMetArgProGlyGlnSerAsnValTyrSerSerTyrGlyGlyS	575
2841	Db	GAAGGTCGAAGAG-----ACTGGGAATAGCGTACCTATAGCTCTTCTGAAAAACC	2891
575	Qy	erThrAspTyrAspSerSerGlnTyrAsnGluAspMetLysLysPheArgLysMetAlaL	595
2892	Db	CGAATGACATC---ACACCAACCGTATGGAAACAATAAGAGAGATTTC-----	2935
595	Qy	euGlyThrGlnGluTyrAsnAlaThrGlyGluTyrSerAsnProThrSerAspTyrGlyL	615
2936	Db	-----GACACAGGTTCAAGCGATGGTTACACTTCAGATATATGGAG	2975
615	Qy	euTyrProSerGlySerSerGluGlyGlnThrThrArgGluMetGluMetGlyLysI	635
2976	Db	TTAACCTTCTCAGTCGAGCAGTGAACATCAACAGGTGAATACTTAGTTCCACAGTTCAA	3035
635	Qy	leiLysArgThrGlyGlnGlyTyrSerGlyProSerLeu	647
3036	Db	TA-----GGCAAGTTTCCACCAATTA	3058

Search completed: May 12, 2004, 22:54:05
Job time : 666 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 12, 2004, 22:25:16 ; Search time 3599 Seconds
(without alignments)
5368.389 Million cell updates/sec

Title: US-10-086-464-2
Perfect score: 3453
Sequence: 1 MSSAPSGTGGPPSPSNT.....REMEMGKIRGTGGYSGPSL 647

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues
Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10086464/runat_11052004_121815_7853/app query.fasta_1.839
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10086464_QCGN_1_1_4237 @runat_11052004_121815_7853 -NCPU=6 -ICPU=3
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hcc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hcc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1092.5	31.6	1084	12	BP184771
2	1090.5	31.6	824	14	CD435142
3	1086	31.5	788	14	CF436437
4	1071	31.0	789	14	CF436655
5	1069	31.0	759	12	BM408099
6	1066	30.9	785	14	CF436122
7	1065	30.8	757	12	BG596561
8	1063	30.8	731	14	CB655196
9	1047	30.3	692	12	BG441204
10	1030	29.8	723	14	CD839231
11	1027.5	29.8	1053	11	AY108241
12	1025	29.7	1016	11	AY108243
13	1018	29.5	666	12	BM358715
14	1013	29.3	695	13	CA072174
15	983	28.5	715	13	BU099573
16	961	27.8	770	13	BQ999193
17	931	27.0	611	14	CA237156
c 18	931	27.0	859	29	CC725849
19	929.5	26.9	809	14	CA765135
20	928	26.9	850	29	CC668754
c 21	923.5	26.7	769	14	CK283399
22	921	26.7	618	14	CA269355
23	921	26.7	651	14	CF478389
24	920	26.6	673	14	CA164704
c 25	919.5	26.6	938	29	CG333846
c 26	917.5	26.6	949	29	CG436431
27	912	26.4	679	13	BQ404121
28	904.5	26.2	674	14	CA298046
29	903	26.2	595	14	CF015663
30	903	26.2	637	13	BQ134241
31	903	26.2	645	13	BQ506869
c 32	897.5	26.0	899	14	CK096498
33	895.5	25.9	682	13	CA095337
34	895	25.9	641	14	CA248316
c 35	895	25.9	725	29	CG450877
c 36	893.5	25.9	803	14	CF243440
37	891	25.8	646	13	BQ240617
38	891	25.8	652	13	BQ849683
39	891	25.8	666	13	BU499754
40	891	25.8	691	14	CA237152
41	889	25.7	626	13	BQ582873
42	888.5	25.7	593	10	BF176907
43	887	25.7	622	14	CA999610
44	887	25.7	673	13	BQ410602
45	886.5	25.7	573	9	AV551753

ALIGNMENTS

RESULT 1
BP184771
LOCUS
DEFINITION BP184771 pNS rice panicle cDNA, germ cell generating stage Oryza sativa (japonica cultivar-group) cDNA, mRNA sequence.
ACCESSION BP184771
VERSION BP184771.1
KEYWORDS EST.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.

```

REFERENCE 1 (bases 1 to 1084)
AUTHORS Moriguchi,K., Ito,Y., Yamazaki,Y. and Kurata,N.
TITLE Finding of various plant nuclear proteins using yeast nuclear
        transportation trap system - a proteomal approach
JOURNAL Unpublished (2003)
COMMENT Contact: Kazuki Moriguchi
        Plant Genetics
        National Institute of Genetics
        Yata 1111, Mishima, Shizuoka 411-8540, Japan
        Tel: 81-55-981-6872
        Fax: 81-55-981-6879
        Email: kmoriguc@lab.nig.ac.jp
        cDNA clone obtained from nuclear transportation trap system
        encoding a protein similar to Oryza sativa (japonica
        cultivar-group) putative receptor protein kinase PERK1.
        Location/Qualifiers
FEATURES
    source
        1..1084
            /organism="Oryza sativa (japonica cultivar-group)"
            /mol_type="mRNA"
            /db_xref="taxon:39947"
            /tissue_type="panicle"
            /dev_stage="germ cell generating stage"
            /clone_lib="pNS rice panicle cDNA, germ cell generating
            stage"

ORIGIN
Alignment Scores:
Pred. No.: 4.98e-26 Length: 1084
Score: 1092.50 Matches: 212
Percent Similarity: 82.88% Conservative: 30
Best Local Similarity: 72.80% Mismatches: 43
Query Match: 31.64% Indels: 7
DB: 12 Gaps: 3

US-10-086-464-2 (1-647) x BP184771 (1-1084)

Qy 360 LeuHisGlyGluGlyArgProThrMetGluTrpSerThrArgLeuLysIleAlaLeuGly 379
Db 2 TTGCAGGAAAGGCCGACCAACATGGAGTGGCCCAACAGACTAAAGATTGCTTTGGGA 61
Qy 380 SerAlaLysGlyLeuSerTyrLeuHisGluAspCysAsnProLysIleIleHisArgAsp 399
Db 62 GCTGCAAGGGTTTAGCTTATCTTCATGAAGACTGCCATCTCTAAGATCATCTCCATCGTAT 121
Qy 400 IleLysAlaSerAsnIleLeuLysAspPheLysPheGluAlaLysValAlaAspPheGly 419
Db 122 ATTAAGCGCTCAACATCTCTTGTGATTTTAAGTTTGAATCTAAGTTGCTGATTTTGA 181
Qy 420 LeuAlaLysIleAlaSerAspThrAsnThrHisValSerThrArgValMetGlyThrPhe 439
Db 182 CTTGCTAAGTTTCCACAGTGATAATAACACTCATGTTTCGACAAGAGTAATGGCAGCTTTT 241
Qy 440 GlyTyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPhe 459
Db 242 GGATATCTAGACACAGAGTAGCATCTCTGGCAAGCTCATCTGAGAAATCAGATGCTCTTC 301
Qy 460 SerPheGlyValValLeuLeuGluLeuLeuThrGlyArgArgProValAlaAsnAsn 479
Db 302 TCCTATGGAGTTATGCTCTTCTTGGTTAATAACTGGTGTGGCCAGTTGTATACAAGTCAA 361
Qy 480 ValTyrValAspAspSerLeuValAspTrpAlaArgProLeuLeuAsnArgAlaSerGlu 499
Db 362 ACATATATGGATGACAGCTTGGTGTGACATGGCCAAAGGCCCTTTACTGATGCAAGCACTTGAG 421
Qy 500 GlnGlyAspPheGluGlyLeuAlaAspAlaLysMetAsnGlyTyrAspArgGluGlu 519
Db 422 AATGGTAACACAGGAGCTTAGTAGTCTCGGCTGGGAAGGATTTCAATCCCAATGAG 481
Qy 520 MetAlaArgMetValAlaCysAlaAlaCysValArgHisSerAlaArgArgPro 539
Db 482 ATGGCGAAGATGATGCTTGTGCAGCTGCATGCGGTACGCCATTCGCTCGTGGTGGCCCA 541
Qy 540 ArgMetSerGlnIleValArgAlaLeuGluGlyAsnValSerLeuSerAspLeuAsnGlu 559

```

```

Db 542 CGCATGACGACAGGTGTCCGGGCTTTTGAAGGTGACGTGCTTTGGAGGATCTTAATGAA 601
Qy 560 GlyMetArgProGlyGlnSerAsnValTy+SerSerTyrGlyGlySerThrAspTyrAsp 579
Db 602 GGTGTTCCGGCTGGTTCACAGCCGCTATTTTGGATCGGTAC---AGCAGCTCTGACTATGAT 658
Qy 580 SerSerGlnTyrAsnGluAspMetLysLysPheArgLysMetAlaLeuGlyThrGlnGlu 599
Db 659 TCTGCCAATACACAGGAGGACATGAAGAAGTTCAGGAAGATGGCTTTT---ACCAACAAT 715
Qy 600 TyrAsnAlaThrGlyGluTyrSerAsnProThrSerAspTyrGlyLeuTyrProSerGly 619
Db 716 AATGATACGAGCAGTCAATACAGCGCACCAACACGAGGATGAGCAGATACCTCTGCA 775
Qy 620 SerSerSerGluGlyGlnThrThrArgGluMetGluMetGlyLysIleLysArgThrGly 639
Db 776 TCRAAGCAGGAGGGCCCAACAAACCCAGGAAGTCGAGAGCAGGACACCAACGAGAGGGCGC 835
Qy 640 -----GlnGlyTyrSerGlyProSer 646
Db 836 TACAGTGGCTACAGCTCAGGATACAGCGGAGCCTCA 871

RESULT 2
CD435142 LOCUS CD435142 824 bp mRNA linear EST 03-JUN-2003
DEFINITION EL01N0355D03.b Endosperm_3 Zea mays cDNA, mRNA sequence.
ACCESSION CD435142
VERSION CD435142.1 GI:31350785
KEYWORDS EST.
SOURCE Zea mays
ORGANISM Zea mays
            1 (bases 1 to 824)
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
            Lai,J., Dey,N., Kim,C.S., Becraft,P., Larkins,B., Linton,E. and
            Messing,J.
            Sequencing of the maize endosperm ESTs
            Unpublished (2002)
            Contact: Lai, Jinsheng
            Dr. Joachim Messing's lab
            Waksman Institute, Rutgers University
            190 Frelinghuysen Rd., Piscataway, NJ 08854, USA
            Tel: 732-445-3801
            Fax: 732-445-5735
            Email: jlai@waksman.rutgers.edu
            Seq primer: T3.
FEATURES
    Location/Qualifiers
    source
        1..824
            /organism="Zea mays"
            /mol_type="mRNA"
            /cultivar="W22"
            /db_xref="taxon:4577"
            /tissue_type="Endosperm of 7-23DAP"
            /clone_lib="Endosperm_3"
            /note="vector: pBluescript SK-; Site_1: EcoRI; Site_2:
            XhoI"
ORIGIN
Alignment Scores:
Pred. No.: 4.54e-26 Length: 824
Score: 1090.50 Matches: 206
Percent Similarity: 84.12% Conservative: 27
Best Local Similarity: 74.37% Mismatches: 41
Query Match: 31.58% Indels: 3
DB: 14 Gaps: 2

US-10-086-464-2 (1-647) x CD435142 (1-824)
Qy 334 LeuValGlyTyrCysIleAlaGlyAlaLysArgLeuLeuValTyrGluPheValProAsn 353
Db 2 TTGTTGGCTATTGCAATTTCTGGAGGCACCTTGCTGCTTGTCTATGATGATTTGTCCCAAT 61

```



```

Qy 354 AenAenLeuGluLeuHisLeuHisGlyGluArgProThrMetGluTrpSerThrArg 373
Db 62 AACACATGGATTCACATTCACGCGAAGGTGCGACCACTATGGAGTGGCTGTCTAGA 121
Qy 374 LeuYsleAlaLeuGlySerAlaGlyLeuSerTyLeuHisGluAspCysAenPro 393
Db 122 TTAAGATCAGTTGGGTGCTGCCAAGGTTAGCTTATCTTCATGAAGCTGCCATCCA 181
Qy 394 LysleileHisArgAspilleLysAlaSerAenleileuLeuAsePheLysPheGluAla 413
Db 182 AAGATCATCCATCGTGCATTAAGGCATCTAACATCTCTTGACTTCCATTTGAAGCT 241
Qy 414 LysValAlaAsePheGlyLeuAlaLysleAlaSerAsePheThrAenThrHisValSerThr 433
Db 242 AAGGTTCGTGATTTGGACTTGCAGTTCCTACTACTGATACACACCCATGTTCCGACA 301
Qy 434 ArgValMetGlyThrPheGlyTyLeuAlaProGluTyAlaAlaSerGlyLysLeuThr 453
Db 302 AGAGTAATGGGCACCTTTGGGTATTTGGCACCTGAGTATGCTCTCTGGCAAGCTAACA 361
Qy 454 GluYsSerAsePheValPheSerPheGlyValValLeuLeuGluLeuLeuThrGlyValArg 473
Db 362 GAAAAATCCGATGATTTCTCGAGTCTATGCTTCTTGAGCTTATTTACTTGGCGCGCA 421
Qy 474 ProValAseAlaAenAenValTyValAsePheSerLeuValAsePheProLeu 493
Db 422 CCAGTTCACACCAACCAACATATATGGATGACAGCTTGGTTGACTGGGCAGGCCATTA 481
Qy 494 LeuAenArgAlaSerGluGlnGlyAspPheGluGlyLeuAlaAsePheAlaLysMetAseAen 513
Db 482 CTGATGCGGACACTTGGAGTGGTGAATGATTTTATGATCTTTAGTGGATCTCTCGGCGGAAG 541
Qy 514 GlyTyAspArgGluGluMetAlaArgMetValAlaCysAlaAlaAlaCysValArgHis 533
Db 542 GACTTCAATCCTAATAGATGGCAAGATGATGACCTGTGCGAGCTGCTGTATGTCAGGCAT 601
Qy 534 SerAlaArgArgArgProArgMetSerGlnleValArgAlaLeuGluGlyAseValSer 553
Db 602 TCTGCAGCTGTCGGCCACGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 661
Qy 554 LeuSerAseLeuAenGluGlyMetArgProGlyGlnSerAenValTySerSerTyGly 573
Db 662 TTGGAGGACCTTAAATGAAGGTGTTTCGGCTGGCCATAGCCGCTTCTTTGGGTCTATAC 718
Qy 574 GlySerThrAsePheSerGlnTyAseGlnTyAseGlnTyAseGlnTyAseGlnTyAseGln 593
Db 719 AGCAGCTCCGATTAACGATTTCTGGCCAGTACACAGGAGACATGAAGAAGTTCAAGAAGATG 778
Qy 594 AlaLeuGlyThrGlnGluTyAseAlaThrGlyGluTySerAenProThr 610
Db 779 GCATTCACACACAC-----TATACGAGCGCCATACAGCGCGCCCAACC 823

RESULT 3
LOCUS CF436437
DEFINITION EST672782 normalized cDNA library of onion Allium cepa cDNA clone
ACACM44, mRNA sequence.
ACCESSION CF436437
VERSION CF436437.1 GI:34459127
KEYWORDS EST.
SOURCE Allium cepa (onion)
ORGANISM Allium cepa

REFERENCE 1 (bases 1 to 788)
AUTHORS Havey, M.J., Cheung, F., Van Aken, S., Utterback, T. and Town, C.D.
TITLE Expressed Sequence Tags from a normalized library of mixed onion
JOURNAL tissues (Allium cepa)
COMMENT Unpublished (2003)
Contact: Havey MJ
Department of Horticulture

```

USDA-ARS and University of Wisconsin
1575 Linden Drive, Madison, WI 53706, USA
Tel: 608-262-1830
Fax: 608-262-4743
Email: mjhavey@facstaff.wisc.edu
TIGR sequence name ACACM44TR. For more information:
http://haveylab.hort.wisc.edu
Seq primer: CAG GAA ACA GCT ATG ACC.

FEATURES

Location/Qualifiers

1..788

/organism="Allium cepa"

/mol_type="mRNA"

/cultivar="Red Creole (bulbs), unknown (callus), Ebano &

Texas Legend (roots)"

/db_xref="taxon:4679"

/clone="ACACM44"

/tissue_type="Callus, roots, and young bulbs"

/clone_lib="normalized cDNA library of onion"

/notes="vector: pCMVSPORT6.1-ccdb (Invitrogen); Site 1:

EcoRV (5'); Site 2: NotI (3'); Equal molar amounts of mRNA

from callus, roots, and young bulbs were combined to

synthesize the library. Normalized to enrich for

low-copy transcripts was performed by proprietary

techniques of Invitrogen."

ORIGIN

Alignment Scores:
Pred. No.: 6,07e-26 Length: 788
Score: 1086.00 Matches: 204
Percent Similarity: 87.79% Conservative: 26
Best Local Similarity: 77.86% Mismatches: 32
Query Match: 31.45% Indels: 0
DB: 14 Gaps: 0

US-10-086-464-2 (1-647) x CF436437 (1-788)

```

Qy 260 LysSerThrPheThrTyLeuGluLeuAlaArgAlaThrAseGlyPheSerGluAlaAse 279
Db 2 AAGAGCACCTTCACATATGAGATTTGGCGATAGCTACAAATGGGTTTCCGACTATAT 61
Qy 280 LeuLeuGlyGlnGlyPheGlyTyValHisGlyValLeuProSerGlyLysGlu 299
Db 62 CTCTCTGGCAAGTGGATTTGGATATGTGCACAAAGAGTACTTCCAAACGGTAAAGAA 121
Qy 300 ValAlaValLysGlnLeuLysValGlySerGlyGlnGlyGluArgGluPheGlnAlaGlu 319
Db 122 GTGGCTTCAAGCAGTGTGAAGCCGGAAGCGGAGCGGAGCGGTGAGTTTCAAGCAGAG 181
Qy 320 ValGluLeileSerArgValHisHisArgHisLeuValSerLeuValGlyTyCysile 339
Db 182 GTTGAATATATCAGTCGATGCATCATAGGCATTTGGTTCTTTAGTTGGCTATTGCATT 241
Qy 340 AlaGlyAlaLysArgLeuLeuValTyGluPheValProAseAenAenLeuGluLeuHis 359
Db 242 TCAGGAGATCATAGATTCGTTCTATGAATATGTTCTTAATAAAACCCCTTGAGTTCCAT 301
Qy 360 LeuHisGlyGluGlyArgProThrMetGluTrpSerThrArgLeuLysleileAlaLeuGly 379
Db 302 TTGCATGAAAAGATGTTCCACTTATGATGGCAACGGGTGAAATTTGCTTTGGGT 361
Qy 380 SerAlaLysGlyLeuSerTyLeuHisGluAspCysAseProLysleileHisArgAse 399
Db 362 TCTGCCAAGGGTTGGCATATCTCCATGAAGATTTGTCATCCCAAAATTTATTCATCGTAT 421
Qy 400 IleLysAlaSerAseNleleuLeuAsePheLysPheGluAlaLysValAlaAsePheGly 419
Db 422 ATTAAGCAGCAAAATATCTCTTGAGATCAACTTTGAGGCTAAGGTTGCAGATTTTGGC 481
Qy 420 LeuAlaLysleileAlaSerAsePheThrAenThrHisValSerThrArgValMetGlyThrPhe 439
Db 482 CTTCGCAAGTTCACCTGTTGAAAATAACACCCATGTTTCTACAGAGTCATGGGGAACATTT 541
Qy 440 GlyTyLeuAlaProGluTyAlaAlaSerGlyLysLeuThrGluLysSerAseValPhe 459

```

```

|||||
542 GGGTATCTGGCACCAGATATGCAATCCCTCGTAAACTAAGTAAATACAGATGCTCTTT 601
Db

460 SerPheGlyValValLeuLeuGluLeuIleThrGlyArgArgProValAspAlaAsnAsn 479
Qy |||||||
Db |||||||

602 TCATTGGAGTCATGCTTTTGGAGCTGATTAAGGAGCGACCAAGTGTGACCAACCAA 661
Qy |||||||
Db |||||||

480 ValTyrValAspSerLeuValAspTrpAlaArgProLeuLeuAsnArgAlaSerGlu 499
Qy |||||||
Db |||||||

662 ACTTATACAGATGATAGTTGGTGCATTTGGGCAAGACCAATGCTGCAGAAAGGCTATGGAG 721
Qy |||||||
Db |||||||

500 GlnGlyAspPheGluGlyLeuAlaAspAlaLysMetAsnAsnGlyTyrAspArgGluGlu 519
Qy |||||||
Db |||||||

722 CATGATGATTATGATGAGCTCATTTGACAGAAGGCTTAGGAGATAAATACCAACATGACGAA 781
Qy |||||||
Db |||||||

520 MetAla 521
782 ATGGCA 787

RESULT 4
CF436655
LOCUS
DEFINITION EST673000 normalized cDNA library of onion Allium cepa cDNA clone
ACCK26, mRNA sequence.
ACCESSION CF436655
VERSION EST.
KEYWORDS Allium cepa (onion)
ORGANISM Allium cepa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Alliaceae;
Allium.
1 (bases 1 to 789)
HAVELY.M.J., Cheung, F., Van Aken, S., Utterback, T. and Town, C.D.
Expressed Sequence Tags from a normalized library of mixed onion
tissues (Allium cepa)
Unpublished (2003)
Contact: Havey MJ
Department of Horticulture
USDA-ARS and University of Wisconsin
1575 Linden Drive, Madison, WI 53706, USA
Tel: 608-262-1830
Fax: 608-262-4743
Email: mjhavey@facstaff.wisc.edu
TIGR sequence name ACCK26T8. For more information:
http://haveylab.hort.wisc.edu
Seq primer: CAG GAA ACA GCT ATG ACC.
Location/Qualifiers
1..789
/organism="Allium cepa"
/mol_type="mRNA"
/cultivar="Red Creole (bulbs), unknown (callus), Ebano &
Texas Legend (roots)"
/db_xref="taxon:4679"
/clones="ACCK26"
/tissue_type="Callus, roots, and young bulbs"
/clone_lib="normalized cDNA library of onion"
/note="Vector: pCMVSPORT6.1-ccdb (Invitrogen); Site 1:
EcoRV (5'); Site 2: NotI (3'); Equal molar amounts of mRNA
from callus, roots, and young bulbs were combined to
synthesize the library. Normalization to enrich for
low-copy transcripts was performed by proprietary
techniques of Invitrogen."

FEATURES
source

ORIGIN
Alignment Scores:
Pred. No.: 1,82e-25 Length: 789
Score: 1071.00 Matches: 203
Percent Similarity: 86.64% Conservative: 24
Best Local Similarity: 77.48% Mismatches: 35
Query Match: 31.02% Indels: 0
DB: 14 Gaps: 0

```

```

US-10-086-464-2 (1-647) x CF436655 (1-789)

Qy 260 LysSerThrPheThrTyrGluGluLeuAlaArgAlaThrAsnGlyPheSerGlyAlaAsn 279
Db |||||||
Qy 280 LeuLeuGlnGlnGlyPheGlyTyrValHisGlyValLeuProSerGlyValGlu 299
Db |||||||
Qy 62 CTTCTTGGGCAAGGTGGATTTGGATATGTGCACAAAGGAGTACTTCCAAAGCTAAAGAA 121
Db |||||||
Qy 300 ValAlaValLysGlnLeuLysValGlySerGlyGlnGlyGluArgGluPheGlnAlaGlu 319
Db |||||||
Qy 122 GTGGCTATCAAGCAGTTGAAAGCCGGAAGCGACAAAGGGAGCGTGTGATTTCAAGCAGAG 181
Qy 320 ValGluIleLeuSerArgValHisHisArgHisLeuValSerLeuValGlyTyrCysIle 339
Db |||||||
Qy 182 GTTGAATTTATCAGTCGAGTGCATCATAGGCATTTGGTTCTTTAGTTGGCTATTGCATT 241
Qy 340 AlaGlyAlaLysArgLeuLeuValTyrGluPheValProAsnAsnAsnLeuLeuHis 359
Db |||||||
Qy 242 TCAGGAGATCATAGATTGCTTGTCTATGAATATGTTCTATAAAACCCCTTGAGTTCCAT 301
Qy 360 LeuHisGlyGluGlyArgProThrMetGluTrpSerThrArgLeuLysIleAlaLeuGly 379
Db |||||||
Qy 302 TTGCATGGAAAGAGATGTTCCACCTATGATGGCCAAACGCGTTAAAAATTGCTTTGGGT 361
Qy 380 SerAlaLysGlyLeuSerTyrLeuHisGluAspCysAsnProLysIleIleHisArgAsp 399
Db |||||||
Qy 362 TCTGCCAAGGCGTTGGCATATCTCCATGAAGATTGTCTATCCCAAAATTTATTCATCGTGT 421
Qy 400 IleLysAlaSerAsnIleLeuIleAspPheLysPheGluAlaLysValAlaAspPheGly 419
Db |||||||
Qy 422 ATTAAGCAGCAAAATATTTCTTTGAGATGAACATTTGAGGCTAAGGTTGCAGATTTTGGC 481
Qy 420 LeuAlaLysIleAlaSerAspThrAsnThrHisValSerThrArgValMetGlyThrPhe 439
Db |||||||
Qy 482 CTTGCAAAAGTTCACTGTTGAAATAACACCCATGTTCTACACGAGTCATCGGAACATTT 541
Qy 440 GlyTyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPhe 459
Db |||||||
Qy 542 GGGTATCTGGCACCAGATATGCAATCTCTGTTAACTAAGTGAATAATCAGATGCTCTTT 601
Qy 460 SerPheGlyValValLeuLeuGluLeuIleThrGlyArgArgProValAspAlaAsnAsn 479
Db |||||||
Qy 602 TCATTGGAGTCATGCTTTTGGAGCTGATTAAGGAGCGACCAAGCTGTGATGCAACCCAA 661
Qy 480 ValTyrValAspSerLeuValAspTrpAlaArgProLeuLeuAsnArgAlaSerGlu 499
Db |||||||
Qy 662 ACTTATACAGATGATAGTTGGTGCATTTGGGCAAGACCAATGCTGCAGAAAGGCTATGGAG 721
Qy 500 GlnGlyAspPheGluGlyLeuAlaAspAlaLysMetAsnAsnGlyTyrAspArgGluGlu 519
Db |||||||
Qy 722 CATGATGATTATGATGAGCTCATTTGACAGAAGGCTTAGGAGATAAATACCAACCATGACGA 781
Qy 520 MetAla 521
782 ATGGCA 787

RESULT 5
BM408099
LOCUS
DEFINITION EST582426 potato roots Solanum tuberosum cDNA clone CPRO33021 5'
end, mRNA sequence.
ACCESSION BM408099
VERSION BM408099
KEYWORDS EST.
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum.
1 (bases 1 to 759)
REFERENCE
1 van der Hoeven, R., Sun, H., Karamycheva, S.A., Tsai, J., Van Aken, S.,

```

Utterback, T., Chiening, A., Bougri, O., Buell, C.R., Ronning, C., Tanksley, S. and Baker, B.
 Generation of ESTs from potato roots
 Unpublished (2001)
 Contact: Robin Buell
 The Institute for Genomic Research
 9712 Medical Center Dr, Rockville, MD 20850, USA
 Email: potato-array@tigr.org
 This clone can be obtained from the University of Arizona Genomics Institute. Orders can be made through URL:
 http://genome.arizona.edu/orders/
 Seq primer: T3.

FEATURES

Location/Qualifiers
 1. .759
 /organism="Solanum tuberosum"
 /mol_type="mRNA"
 /cultivar="Kennebec"
 /db_xref="taxon:4113"
 /clone="cPRO33021"
 /tissue_type="roots"
 /dev_stage="in vitro grown stem cuttings"
 /lab_host="SOLR"
 /clone_lib="potato roots"
 /notes="vector: pBluescript SK(-); Site 1: EcoRI; Site 2: XhoI; supplier: Cornell University, Tanksley Lab; sequencing: The Institute for Genomic Research. Roots were isolated from in vitro grown stem cuttings on CM medium. Roots were isolated two weeks after placing the stem cuttings from in vitro grown plants on medium."

ORIGIN

Alignment Scores:
 Pred. No.: 2.04e-25 Length: 759
 Score: 1069.00 Matches: 200
 Percent Similarity: 89.29% Conservative: 25
 Best Local Similarity: 79.37% Mismatches: 27
 Query Match: 30.96% Indels: 0
 DB: 12 Gaps: 0

US-10-086-464-2 (1-647) x BM408099 (1-759)

312 GlyGluArgGluPheGlnAlaGluValGluIleSerArgValHisHisArgHisLeu 331
 DB 2 GGGGAACGTGAATTCAGGCGGAAGTTGAGATTATTAGCGAGTACATCAAGCATCTT 61
 332 ValSerLeuValGlyTyrCysIleAlaGlyAlaLysArgLeuValTyrGluPheVal 351
 DB 62 GTGTCTCTTGTGGTACTGCAATTAAGGCTCAGAGACTGCTTGTATGAGTTTGT 121
 352 ProAsnAsnLeuGluLeuHisLeuHisGlyGluArgProThrMetGluTrpSer 371
 DB 122 CCAACAATATCTTGGAAATTCATTACAGGAAGGAAGGAGCTCTTTGGATTGGCCA 181
 372 ThrArgLeuLysIleAlaLeuGlySerAlaLysGlyLeuSerTyrLeuHisGluAspCys 391
 DB 182 ATACGGCTAAAGATTGCTTAGGCTCAGCTAAAGGAGCTGGCATATCTGCATGAAGACTGC 241
 392 AsnProLysIleLeHisArgAspIleLysAlaSerAsnIleLeuIleAspPheLysPhe 411
 DB 242 CAACCGAAATCATTCACCGGTGATATCAAGGAGCTAATATATCTATCGACTTTAAATTT 301
 412 GluAlaLysValAlaAspPheGlyLeuAlaLysIleAlaSerAspThrAsnThrHisVal 431
 DB 302 GAGCTAAGGTGCTGATTTGGACTTGGCAAGCTAACTCTCTGATGTTAATCTCATGTC 361
 432 SerThrArgValMetGlyThrPheGlyTyrLeuAlaProGluTyrAlaAlaSerGlyLys 451
 DB 362 TCCACCAAGATGATGGGAACCTTTGGGTATTGGCTCCAGATATGCTTCTCTGGGAAG 421
 452 LeuThrGluLysSerAspValPheSerPheGlyValValLeuLeuGluLeuIleThrGly 471
 DB 422 CTTACAGACAAGTCAGACGATTTCTCTTTGGTGAATGCTTCTTGATGTTGATAACTGGA 481

472 ArgArgProValAspAlaAsnValTyrValAspAspSerLeuValAspTrpAlaArg 491
 DB CGTGGGCTGTGACTCTACTCAATCATCATCAAGATAGTTTGGTGGAGCTGGGACGT 541
 492 ProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGluGlyLeuAlaAspAlaLysMet 511
 DB 542 CCATTACTCACAGAGCTTTAGAAAGATGAAAGTTTGTATACCTTGTGTGTCGCGCTA 601
 512 AsnAsnGlyTyrAspArgGluGluMetAlaArgMetValAlaCysAlaAlaCysVal 531
 DB 602 GAAATGATTATACCATATAGATGGCTCGCATGGTTGCTTGTGCTGCTGCTGTG 661
 532 ArgHisSerAlaArgArgProArgMetSerGlnIleValArgAlaLeuGluGlyAsn 551
 DB 662 CGTCATTCAGAAAGCGTAGACCAAGATGACACAGGTTCTCCGAGCTTGGAAAGGAGAT 721
 552 ValSerLeuSerAspLeuAsnGluGlyMetArgPro 563
 DB 722 GTCTCATTTACAGACCTTTACGAAGGGATTAAACCT 757

RESULT 6

CF436122 785 bp mRNA linear EST 04-SEP-2003
 LOCUS EST672467 normalized cDNA library of onion Allium cepa cDNA clone
 DEFINITION ACACJ26, mRNA sequence.

ACCESSION CF436122

VERSION CF436122.1 GI:34458812

KEYWORDS EST.

SOURCE Allium cepa (onion)

ORGANISM

REFERENCE 1 (bases 1 to 785)
 AUTHORS Havey, M.J., Cheung, F., Van Aken, S., Utterback, T. and Town, C.D.
 TITLE Expressed Sequence Tags from a normalized library of mixed onion tissues (Allium cepa)

JOURNAL Unpublished (2003)

COMMENT Contact: Havey, M.J.
 Department of Horticulture
 USDA-ARS and University of Wisconsin
 1575 Linden Drive, Madison, WI 53706, USA
 Tel: 608-262-1830
 Fax: 608-262-4743
 Email: mjhavey@facstaff.wisc.edu
 TIGR sequence name ACACJ26TR. For more information:
 http://haveylab.hort.wisc.edu
 Seq primer: CAG GAA ACA GCT ATG ACC.

FEATURES

source
 1. .785
 /organism="Allium cepa"
 /mol_type="mRNA"
 /cultivar="Red Creole (bulbs), unknown (callus), Ebano & Texas Legend (roots)"
 /db_xref="taxon:4679"
 /clone="ACACJ26"
 /tissue_type="Callus, roots, and young bulbs"
 /clone_lib="normalized cDNA library of onion"
 /notes="Vector: pCMVSPORT6.1-ccdb (Invitrogen); Site 1: EcoRV (5'); Site 2: NotI (3'); Equal molar amounts of mRNA from callus, roots, and young bulbs were combined to synthesize the library. Normalization to enrich for low-copy transcripts was performed by proprietary techniques of Invitrogen."

ORIGIN

Alignment Scores:
 Pred. No.: 2.61e-25 Length: 785
 Score: 1066.00 Matches: 204
 Percent Similarity: 86.64% Conservative: 23
 Best Local Similarity: 77.86% Mismatches: 34
 Query Match: 30.87% Indels: 1
 DB: 14 Gaps: 0

```
US-10-086-464-2 (1-647) x CP436122 (1-785)
Qy 260 LysSerThrPheThrTyrGluGluLeuAlaArgAlaThrAsnGlyPheSerGluAlaAsn 279
Db 2 AAGAGCACTTTCACATATGAAGATTGGCGATAGCTACAAATGGTTTCCGACTATAAT 61
Qy 280 LeuLeuGlyGlnGlyGlyPheGlyTyrValHisLysGlyValLeuProSerGlyLysGlu 299
Db 62 CTTCTTGGCGAAGGTGGATTGGATATGTGCACAAAGGAGTACTTCCAAACGGGTAAAGAA 121
Qy 300 ValAlaValLysGlnLeuLysValGlySerGlyGlnGlyGluArgGluPheGlnAlaGlu 319
Db 122 GTGGCTATCAAGCAGTGTGAAGCCGGAAGCGGACAGGGGAGCGTGAGTTTCAAGCAGAG 181
Qy 320 ValGluLeuLeuSerArgValHisHisArgHisLeuValSerLeuValGlyTyrCysIle 339
Db 182 GTTGAATTTATCAGTCGAGTCATCATAGGCATTTGGTTCTTTAGTTGGCTATTGCATT 241
Qy 340 AlaGlyAlaLysArgLeuLeuValTyrGluPheValProAsnAsnLeuGluLeuHis 359
Db 242 TCAGGAGATCATAGATTGCTTGTCTATGAATATGTTCTTAATAAACCCTTGAGTTCCAT 301
Qy 360 LeuHisGlyGluGlyArgProThrMetGluTrpSerThrArgLeuLysIleAlaLeuGly 379
Db 302 TTGCATGGAAAGATGTTCCACCTATGGATTGGCCACCGGTTAAAAATTCCTTTGGGT 361
Qy 380 SerAlaLysGlyLeuSerTyrLeuHisGluAspCysAsnProLysIleIleHisArgAsp 399
Db 362 TCTGCCAAGGGTTGGCATATCTCCATGAAGATTGTCTATCCCAAAATTTATTCATCGTAT 421
Qy 400 IleLysAlaSerAsnIleLeuIleAspPheLysPheGluAlaLysValAlaAspPheGly 419
Db 422 ATTAAGCAGCAATATTTCTTTCAGATGAACATTTTCAGGCTAAGGTTGCAGATTTTGGC 481
Qy 420 LeuAlaLysIleAlaSerAspThrAsnThrHisValSerThrArgValMetGlyThrPhe 439
Db 482 CTTCCAAAGTTCACCTGTTGAAATAACACCATGTTTCTACACGAGTCATGGGAACATTT 541
Qy 440 GlyTyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPhe 459
Db 542 GGGTATCTGGCACCAGAAATATGATCTCTCTGTTAACTAAGTATGAATAATCAGATGTCTTT 601
Qy 460 SerPheGlyValValLeuLeuGluLeuIleThrGlyArgArgProValAspAlaAsnAsn 479
Db 602 TCATTTCAGTGCATGCTTTTGGAGCTGATTACTTGGAAAGCGACCATGTTGATGCAACCCAA 661
Qy 480 ValTyrValAspAspSer-LeuValAspTrpAlaArgProLeuLeuAsnArgAlaSerGlu 499
Db 662 ACTTATACAGATGATAGTTTGGTTCGATTGGGCAAGACCATTTCTGCAGAAAGGCTATGGA 721
Qy 499 uGlnGlyAspPheGluGlyLeuAlaAspAlaLysMetAsnAsnGlyTyrAspArgGluGlu 519
Db 722 GCATGATGATTATCATGAGCTCATTTGACAGAAAGGCTAGGAGATAATAACAACCATGACA 781
Qy 519 uMet 520
Db 782 AATG 785

RESULT 7
LOCUS BG596561
DEFINITION EST495239 cSTS Solanum tuberosum cDNA clone cSTS15A23 5' sequence, mRNA sequence.
ACCESSION BG596561
VERSION BG596561.1 GI:13614701
KEYWORDS EST.
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliopsida; eudicotyledons; core eudicots;
asterids; lamiales; Solanales; Solanaceae; Solanum.
REFERENCE 1 (bases 1 to 757)

AUTHORS van der Hoeven, R., Bezzerides, J., Sun, H., Cho, J., Chiemingo, A.,
Bougri, O., Buell, C.R., Ronning, C., Tanksley, S. and Baker, B.
TITLE Generations of ESTs from sprouting potato eyes
JOURNAL Unpublished (2000)
COMMENT Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Email: potato-array@tigr.org
This clone can be obtained from the University of Arizona Genomics
Institute. Orders can be made through URL:
http://genome.arizona.edu/orders/
Seq primer: M13P-R.
FEATURES
source
1..757
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="cSTS15A23"
/tissue_type="sprouting eyes from tubers"
/dev_stage="12-14 weeks post harvest"
/lab_host="SOLR"
/clone_lib="cSTS"
/notes="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; Various sizes of sprouting eyes (2mm to 15mm) were
taken from tubers. The tubers were incubated at 26C in the
dark for 2-3 weeks prior to sprouting. The eyes were
frozen in liquid nitrogen immediately upon removal from
tubers."
ORIGIN
Alignment Scores: 2.72e-25 Length: 757
Pred. No.: 1065.00 Matches: 200
Score: 89.16% Conservative: 22
Percent Similarity: 80.32% Mismatches: 27
Best Local Similarity: 30.84% Indels: 0
Query Match: 12 Gaps: 0
DB: 0
US-10-086-464-2 (1-647) x BG596561 (1-757)
Qy 301 AlaValLysGlnLeuLysValGlySerGlyGlnGlyGluArgGluPheGlnAlaGluVal 320
Db 1 GCAGTTAAACAGCTTAAGGCTGGAGTGGACAGGGGACGTAATTCAGCGGAGATT 60
Qy 321 GluIleIleSerArgValHisHisArgHisLeuValSerLeuValGlyTyrCysIleAla 340
Db 61 GAGATTATTAGCCGAGTACATCACAGCATCTTGTGTCTCTTGTGGATACTGCAATTACT 120
Qy 341 GlyAlaLysArgLeuLeuValTyrGluPheValProAsnAsnAsnLeuGluLeuHisLeu 360
Db 121 GGGGCTCAGAGCTGCTTGTATTAGTTTGTTCACAAACCAATACTTTTGGAAATTTCAATTA 180
Qy 361 HisGlyGluGlyArgProThrMetGluTrpSerThrArgLeuLysIleAlaLeuGlySer 380
Db 181 CACGAAAGGGAAGCCCTCTTTGGATTGGCCATACGGCTAAAGATGTCTTAGGGTCA 240
Qy 381 AlalysGlyLeuSerTyrLeuHisGluAspCysAsnProLysIleIleHisArgAspIle 400
Db 241 GCTAAAGGACTGGCATATCTGCATGAAGACTGCCAACCGAAATCAATTCACCGTGATATC 300
Qy 401 LysAlaSerAsnIleLeuIleAspPheLysPheGluAlaLysValAlaAspPheGlyLeu 420
Db 301 AAGGCACATAATATCTATTCGACTTTAAATTTTTCAGGCTAAGGTTGCTGATTTTGGACTT 360
Qy 421 AlalysIleAlaSerAspThrAsnThrHisValSerThrArgValMetGlyThrPheGly 440
Db 361 GCCAAGCTAACTTCTGATGTTTAATCTCATCTCTCCACAGAGTGATGGGAACCTTTGGG 420
Qy 441 TyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPheSer 460
Db 421 TATTGGCTCCAGAAATATGCTTCTCTGGAAAGCTTACAGACAAAGTCAGAGCTATTCTCC 480
```



```
JOURNAL
COMMENT
of the cotton fiber
Unpublished (2000)
Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Seq primer: TAATACGACTCACTATAGGG
High quality sequence stop: 690.

FEATURES
source
1..692
/organism="Gossypium arboreum"
/mol_type="mRNA"
/strain="AKA"
/cultivar="8400"
/db_xref="taxon:29729"
/clone="GA_Ea0012C15f"
/tissue_type="Fibers isolated from bolls harvested 7-10
dpa"
/lab_host="E. coli"
/clone_libs="Gossypium arboreum 7-10 dpa fiber library"
/note="Vector: pBK-CMV; Site_1: EcoRI; Site_2: XhoI"

ORIGIN
Alignment Scores:
Pred. No.: 9.38e-25 Length: 692
Score: 1047.00 Matches: 198
Percent Similarity: 94.71% Conservat: 17
Best Local Similarity: 87.22% Mismatches: 12
Query Match: 30.32% Indels: 0
DB: 12 Gaps: 0

US-10-086-464-2 (1-647) x BG441204 (1-692)

Qy 247 LeuProProSerProGlyLeuValLeuGlyPheSerLysSerThrPheThrGlu 266
Db 11 CTTCCGCTGGTCCACCTGGTATTTCTTAGTTCTCGAAGACATTTAGCTATGAA 70
Qy 267 GluLeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyPhe 286
Db 71 GAATTAGCGAGCAACGAGTGGCTTCCTCGGAAGTTAACTTCTTGGACAAGTGGTTT 130
Qy 287 GlyTyrValHisLysGlyValLeuProSerGlyLysGluValAlaValLysGlnLeuLys 306
Db 131 GGGTACGTACACAAGAGGATTTCTCCCTAATGGGAAGAGTAGCAGTAAAGCAACTCAAG 190
Qy 307 ValGlySerGlyGlnGlyArgGluPheGlnAlaGluValGluLeuLeuSerArgVal 326
Db 191 GCTGGAATGGCGAAGCGGAGAGAGAAATTTGAGGCTGAAGTTGAGATCATTTAGCCGCTC 250
Qy 327 HisHisArgHisLeuValSerLeuValGlyTyrCysIleAlaGlyAlaLysArgLeuLeu 346
Db 251 CATCACAAACATCTGCTCTCATTTGTCGATCTGATTTCTGGGACAATTAAGATGCTT 310
Qy 347 ValTyrGluPheValProAsnAsnLeuGluLeuHisLeuHisGlyGluGlyArgPro 366
Db 311 GTTTATGAGTTTGTTCACCAACAACACCTTGGAGTTTCACTTGCATGGGAGGCGGCGCTG 370
Qy 367 ThrMetGluTyrSerThrArgLeuLysIleAlaLeuGlySerAlaLysGlyLeuSerTyr 386
Db 371 ACCATGGAATGGCGAAGGATGAAATTTGCTTTAGGATCTGCAAAAGGACTGGCATAT 430
Qy 387 LeuHisGluAspCysAsnProLysIleIleHisArgAspIleLysAlaSerAsnIleLeu 406
Db 431 CTTTCATGAAGATTGTCATCCTAAGATCATTCACCGTGATTAAGCGCGCTAATATCTG 490
Qy 407 IleAspPheLysPheGluAlaLysValAlaAspPheGlyLeuAlaLysIleAlaSerAsp 426
Db 491 TTGGATTTCAAGTTTGAAGCAAGAGTTCCTGATTTTGGACTAGCGAAATTTGCTTCGAT 550
Qy 427 ThrAsnThrHisValSerThrArgValMetGlyThrPheGlyTyrLeuAlaProGluTyr 446

Db 551 GTCAACACGACGCTCTCCACCAGGGTGATGGTACTTTTCGGGTATTTTAGCCCTGAGTAT 610
Qy 447 AlaAlaSerGlyLysLeuThrGluLysSerAspValPheSerPheGlyValValLeuLeu 466
Db 611 GCTTCAAGTGAAGCTCACTGATAATCAATGATGTTTCTCCTTCGGGGTCATGCTTTTG 670
Qy 467 GluLeuLeuThrGlyArgArg 473
Db 671 GAGTTGATTACCGGTACAGAG 691

RESULT 10
CD839231
LOCUS CD839231 723 bp mRNA linear EST 10-JUL-2003
DEFINITION RFO2.114104F010529 RFO2 Brassica napus cDNA clone RFO2114104, mRNA
sequence.
ACCESSION CD839231
VERSION CD839231.1 GI:32521171
KEYWORDS EST.
SOURCE Brassica napus (rape)
ORGANISM Brassica napus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE
1 (bases 1 to 723)
Genopiante.
Genopiante, a major partnership french program in plant genomics
Unpublished (2003)
Contact: Genopiante
Genopiante
93, rue Henri Rochefort 91025 EVRY CEDEX France
Tel: 33 1 69 47 54 00
Fax: 33 1 69 47 54 10
This sequence has been generated in the framework of the french
plant genomics programme 'Genopiante' (http://www.genopiante.com
and http://genopiante-info.infobiogen.fr).

FEATURES
source
1..723
/organism="Brassica napus"
/mol_type="mRNA"
/cultivar="samourai (restored line)"
/db_xref="taxon:3708"
/clone="RFO2114104"
/tissue_type="anthers"
/clone_lib="RFO2"

ORIGIN
Alignment Scores:
Pred. No.: 3.37e-24 Length: 723
Score: 1030.00 Matches: 199
Percent Similarity: 100.00% Conservat: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 29.83% Indels: 0
DB: 14 Gaps: 0

US-10-086-464-2 (1-647) x CD839231 (1-723)

Qy 449 SerGlyLysLeuThrGluLysSerAspValPheSerPheGlyValValLeuLeuLeu 468
Db 2 AGCGGAAAGCTCACCGAGAAGTCTGACGTTTCTCATTTGGCGTTGTGCTTTGGAGCTC 61
Qy 469 IleThrGlyArgArgProValAspAlaAsnValTyrValAspAspSerLeuValAsp 488
Db 62 ATTACCGGGCGTCCACCGCTTGATGCCAACAAATGCTATGTAGATGACAGCTTAGTTGAC 121
Qy 489 TrpAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGlyLeuAlaAsp 508
Db 122 TGGGACGACCATTTGCTTAACCGAGCATCTGAGCAAGGAGACTTTGAGGGTTTACCTGAT 181
Qy 509 AlaLysMetAsnAsnGlyTyrAspArgGluMetAlaArgMetValAlaCysAlaAla 528
Db 182 GCAAGATGAATAATGGGTATGACAGAGGAGATGGCTCGCATGGTTGCTTGTGCTCGC 241
```

Qy 529 AlaCysValArgHisSerAlaArgArgProArgMetSerGlnIleValArgAlaLeu 548
 Db 242 GCTTGTGTTGCCATTGAGCTCGCGCAGACCTCGCATGAGCCAGATTGTGCGTGCCTTA 301
 Qy 549 GluGlyAsnValSerLeuSerAspLeuAsnGluGlyMetArgProGlyGlnSerAsnVal 568
 Db 302 GAAGGAAATGTATCAGTGTTCAGATCTTAACCAAGGAGATGAGACGAGTCAAGCAATGTA 361
 Qy 569 TyrSerSerTyrGlyGlySerThrAspTyrAspSerSerGlnTyrAsnGluAspMetLys 588
 Db 362 TACAGCTCATACGAGGAGGACCGATATGACTCGAGCCAGTACATGAGACATGAG 421
 Qy 589 LysPheArgLysMetAlaLeuGlyThrGlnGluTyrAsnAlaThrGlyGluTyrSerAsn 608
 Db 422 AAGTTTAGGAAATGGCTCTTGGAACTCAAGAGTACAAACGCCACGGGTGAGTACAGTAAT 481
 Qy 609 ProThrSerAspTyrGlyLeuTyrProSerGlySerSerGlyGlnGlyGlnThrThrArg 628
 Db 482 CCGACCAAGCGACTATGACTGTATCCCGTCTGGTTTCAAGCAGCGAGGCGCAACACACGC 541
 Qy 629 GluMetGluMetGlyLysIleLysArgThrGlyGlnGlyTyrSerGlyProSerLeu 647
 Db 542 GAAATGGAGATGGGGAAGATTAAAGAAACCGGTGAGGGTTATAGTGGACCTTCTCTT 598

RESULT 11

LOCUS AY108241 1053 bp mRNA linear HTC 16-OCT-2002
 DEFINITION Zea mays PC0134818 mRNA sequence.
 ACCESSION AY108241
 VERSION AY108241.1 GI:21211319
 KEYWORDS HTC.

SOURCE

ORGANISM Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.

REFERENCE

1 (bases 1 to 1053)
 Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
 Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
 Maize Mapping Project/DuPont Consensus Sequences for Design of
 Overgo Probes

JOURNAL

Unpublished (2002)

REFERENCE

2 (bases 1 to 1053)

AUTHORS

Coe,E.H.

TITLE

Direct Submission

JOURNAL

Submitted (25-APR-2002) Maize Mapping Project, University of
 Missouri, Columbia, MO 65211, USA
 If you are interested in getting corresponding physical clones,
 these are publicly available from ZmDB and may be found by BLAST
 searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
 www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
 maize cDNA sequences is either Virginia Walbot, Stanford or Pat
 Schnable, Iowa State, then clones may be requested from ZmDB:
 www.zmdb.iastate.edu.

FEATURES

source

1..1053
 Location/Qualifiers
 /organism="Zea mays"
 /mol_type="mRNA"
 /db_xref="MaizeDB:637889"
 /db_xref="taxon:4577"
 /clone_lib="Maize Mapping Project/DuPont Consensus
 Library"

/note="this sequence is part of a project of EST
 assemblies resulting from the application of public
 contigs to seed Dupont contigs; this resource was
 assembled by Dupont as part of a collaboration for the
 overgo addressing of BACs in conjunction with the Maize
 Mapping Project"

ORIGIN

Alignment Scores: 5.61e-24 Length: 1053
 Pred. No.: 1027.50 Matches: 203
 RESULT 12
 AY108243
 LOCUS AY108243 1016 bp mRNA linear HTC 16-OCT-2002
 DEFINITION Zea mays PC0134814 mRNA sequence.
 ACCESSION AY108243
 VERSION AY108243.1 GI:21211321

Percent Similarity: 79.93% Conservative: 28
 Best Local Similarity: 70.24% Mismatches: 46
 Query Match: 29.76% Indels: 12
 Db: 11 Gaps: 4
 US-10-086-464-2 (1-647) x AY108241 (1-1053)
 Qy 365 ArgProThrMetGluTrpSerThrArgLeuLysIle-AlaLeuGlySerAlaLysGlyLe 384
 Db 6 CGTCCG-----GAGTGGCCTGTAGATTAAGATCATGTTTGGGTGCTGCTCAAGGGTTT 59
 Qy 384 uSerTyrLeuHisGluAspCysAsnProLysIleIleHisArgAspIleLysAlaSerAs 404
 Db 60 AGCTTATCTTCATGAAGACTGCCATCCAAAGATCATCCATCGTGACATAAAGGATCTAA 119
 Qy 404 nIleLeuIleAspPheLysPheGluAlaLysValAlaAspPheGlyLeuAlaLysIleAl 424
 Db 120 CATCTCTCTTGACTTCCATTTGAAGCTAAGTTGCTGATTTTGGATCTTGCAGAGTTTCC 179
 Qy 424 aSerAspThrAsnThrHisValSerThrArgValMetGlyThrPheGlyTyrLeuAlaPr 444
 Db 180 TACTGATAACAACACACCCATGTTTCGACAAGAGTAATGGGCACCTTTGGGTATTGGCACC 239
 Qy 444 oGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPheSerPheGlyValVa 464
 Db 240 TGAGTATGCATCTCTTGGCAAGCTTAAACAGAAAAATCCGATGTATTTCTTCGGAGTCAT 299
 Qy 464 lLeuLeuGluLeuIleThrGlyArgArgProValAspAlaAsnValTyrValAspAs 484
 Db 300 GCTTCTTGAGCTTATTACTGGGGCGGACCATGTTGACACCAACCATATATATATGATGA 359
 Qy 484 pSerLeuValAspTrpAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGl 504
 Db 360 CAGCTTGGTTGACTGGGCAAGGCCATTTACTGATCGGAGCCTTTGAGGATGGTGAATATGA 419
 Qy 504 uGlyLeuAlaAspAlaLysMetAsnGlyTyrAspArgGluGluMetAlaArgMetVa 524
 Db 420 TGCTTTAGTGCATCTCGCTGGGAAAGGACTTCAATCTTAATGAGATGGCAAGATGAT 479
 Qy 524 lAlaCysAlaAlaAlaCysValArgHisSerAlaArgArgProArgProArgMetSerGlnI 544
 Db 480 AGCCTGTGCAGCTGATGTAGCCCATCTCGACGTCGTCGCCACGATATGATGATCAGGT 539
 Qy 544 eValArgAlaLeuGluGlyAsnValSerLeuSerAspLeuAsnGluGlyMetArgProGl 564
 Db 540 CGTTCGGGCTTTGGAGGGCAATGTCTTTGGAGGACCTTAATGAAGGTGTTCGSCCTGG 599
 Qy 564 yGlnSerAsnValTyrSerSerTyrGlyGlySerThrAspTyrAspSerSerGlnTyrAs 584
 Db 600 CCATAGCCGGCTTCTTTGGGTCTATAC---AGCAGCTCCGATTACGATTTCTGCCAGTACA 656
 Qy 584 nGluAspMetLysLysPheArgLysMetAlaLeuGlyThrGlnGluTyrAsnAlaThrGl 604
 Db 657 CGAGGACATGAAGAAGTTCAGAAGATGGCATTAACAACAAC-----TATACACGAG 710
 Qy 604 yGluTyrSerAsnProThrSerAspTyrGlyLeuTyrProSerGlySerSerGluGl 624
 Db 711 CCATACAGCGCGCCCAACCATGATATGACAGATACCGTCTGCATCAACGACGAGGG 770
 Qy 624 yGlnThrThrArgGluMetGluMetGlyLysIleLysArgThrGly----- 639
 Db 771 CCACACAGCGCAAGAGATGGAGTCGGGTGCAATGAAGAAGGTGCTTACAGTGTGGCTA 830
 Qy 640 ----GlnGlyTyrSerGlyProSer 646
 Db 831 CAGCTCAGGATACAGCGGAGCCTCG 855

```

KEYWORDS      HTC.
SOURCE         Zea mays
ORGANISM       Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
               Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
               clade; Panicoideae; Andropogoneae; Zea.
REFERENCE      1 (bases 1 to 1016)
AUTHORS        Hainey,C.P., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
               Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
               Zea Mapping Project/DuPont Consensus Sequences for Design of
               Overgo Probes
TITLE          Unpublished (2002)
JOURNAL        2 (bases 1 to 1016)
REFERENCE      Coe,E.H.
AUTHORS        Direct Submission
TITLE          Submitted (25-APR-2002) Maize Mapping Project, University of
               Missouri, Columbia, MO 65211, USA
JOURNAL        If you are interested in getting corresponding physical clones,
               these are publicly available from ZmDB and may be found by BLAST
               searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
               www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
               maize cDNA sequences is either Virginia Walbot, Stanford or Pat
               Schnable, Iowa State, then clones may be requested from ZmDB:
               www.zmdb.iastate.edu.
FEATURES       Location/Qualifiers
               1..1016
               /organism="Zea mays"
               /mol_type="mRNA"
               /db_xref="taxon:4577"
               /db_xref="MaizeDB:637888"
               /clone_lib="Maize Mapping Project/DuPont Consensus
               Library"
               /note="this sequence is part of a project of EST
               assemblies resulting from the application of public
               contigs to seed DuPont contigs; this resource was
               assembled by DuPont as part of a collaboration for the
               overgo addressing of BACs in conjunction with the Maize
               Mapping Project"
ORIGIN
Alignment Scores:
Pred. No.:      6,53e-24      Length:      1016
Score:          1025.00      Matches:    200
Percent Similarity: 82.14%      Conservative: 30
Best Local Similarity: 71.43%      Mismatches: 42
Query Match:     29.68%      Indels:     8
DB:              11          Gaps:        4

US-10-086-464-2 (1-647) x AV108243 (1-1016)
Qy      373 ArgLeuYsIleAlaLeuGlySerAlaLysGlyLeuSerTyrLeuHisGluAspCysAsn 392
Db      3 AGATTAAAGATCGCTCTGGGTGCTGCCAAGGGTTTAGCTTATCTTCATGAAGCTGCCAT 62
Qy      393 ProlySleIleHisArgAspIleLysAlaSerAsnIleLeuIleAspPheLysPheGlu 412
Db      63 CCAAAGATCATCCACCGGCATTAAGGCATCTAACATCTTCTTGACTTCATTAATTTGAA 122
Qy      413 AlaLysValAlaAspPheGlyLeuAlaLysIleAlaSerAspThrAsnThrHisValSer 432
Db      123 GCTATGTTGCTGACTTTGGACTTGGCAAGTTCACCTACTGATATACACACCACCTGTGTCA 182
Qy      433 ThrArgValMetGlyThrPheGlyTyrLeuAlaProGluTyrAlaAlaSerGlyLysLeu 452
Db      183 ACAAGAGTAATGGGCACCTTTGGGTATTTGGCCACCGAGTAGTCAGCATCTGGCAAGCTC 242
Qy      453 ThrGluLysSerAspValPheSerPheGlyValValLeuLeuGluLeuIleThrGlyArg 472
Db      243 ACAGAAAATCCGATGTATTTCTTTCGGAGTCATGCTTCTTGAGCTTATTACTGGGCGG 302
Qy      473 ArgProValAspAlaAsnValTyrValAspAspSerLeuValAspTyrAlaArgPro 492
Db      303 CGACCAATTGACACACCAACCAACATATATGATGACAGCTTGGTTGACTGGGCAAGGCCA 362

```

```

Qy      493 LeuLeuAsnArgAlaSerGluGlnGlyAspPheGluGlyLeuAlaAspAlaLysMetAsn 512
Db      363 TTACTGATGAGAGCGCTCGAGGATGGTGAATATATGCTTTGGTGGATCTCTCGGCTCGGA 422
Qy      513 AsnGlyTyrAspArgGluGluMetAlaArgMetValAlaCysAlaAlaCysValArg 532
Db      423 AAGGACTTCAATCTTAAGAGATGGCAAGATAGCCTGTGCGGCTCTCGGAGCGGACGTG 482
Qy      533 HisSerAlaArgArgArgProArgMetSerGlnIleValArgAlaLeuGluGlyAsnVal 552
Db      483 CATTCGCGACGTCGTGGCCACCAAGTATGATGATGATGATGATGATGATGATGATGATG 542
Qy      553 SerLeuSerAspLeuAsnGluGlyMetArgProGlyGlnSerAsnValTyrSerSerTyr 572
Db      543 TCTTTGGAGGACCTTAATGAAGGTGTTGACCTCGCCATAGCCGCTTCTTTGGGTGCATAT 602
Qy      573 GlyGlySerThrAspTyrAspSerSerGlnTyrAsnGluAspMetLysLysPheArgLys 592
Db      603 ---AGCAGCTCTGATTACGATTCGGGCGCATGTAACGAGGACATGCGAAGTTGAGGAG 659
Qy      593 MetAlaLeuGlyThrGlnGluTyrAsnAlaThrGlyGluTyrSerAsnProThrSerAsp 612
Db      660 ACGCGTTTCAATACAAACACCTACACCCAGCAGC---CAATACAGCGCGCCCAACAGTGAA 716
Qy      613 TyrGlyLeuTyrProSerGlySerSerGluGlyGlnThrThrArgGluMet---Glu 631
Db      717 TACGCCCAGGTGCGGCTCTGGATCAAGCAGCGAGGCGGCCACGACGACGAGGATGGAGG 776
Qy      632 MetGlyLysIleLysArgThrGly-----GlnGlyTyrSerGlyProSer 646
Db      777 TCGGGTCGATGAAGAAAGGTGGCTACAGCGGCTACAGCTCCGATACGCGGAGCCTCG 836

RESULT 13
BM358715
LOCUS      666 bp      mRNA      linear      EST 09-JAN-2002
DEFINITION GA_Ba0012D16r Gossypium arboreum 7-10 dpa fiber library Gossypium
               arboreum cDNA clone GA_Ba0012D16r, mRNA sequence.
ACCESSION   BM358715
VERSION     BM358715.1
KEYWORDS    GI:18099461
SOURCE      Gossypium arboreum
ORGANISM    Gossypium arboreum
REFERENCE    1 (bases 1 to 666)
AUTHORS      Wing,R.A., Frisch,D., Yu,Y., Main,D., Rambo,T., Simmons,J.,
               Henry,D., Wood,T.C., Leslie,A. and Wilkins,T.A.
               An integrated analysis of the genetics, development, and evolution
               of the cotton fiber
TITLE        Unpublished (2000)
JOURNAL      Contact: Wing RA
COMMENT      Clemson University Genomics Institute
               Clemson University
               100 Jordan Hall, Clemson, SC 29634, USA
               Tel: 864 656 7288
               Fax: 864 656 4293
               Email: rwing@clemson.edu
               Total High Quality bases = 646
               Seq primer: TAATACGACTCACTATAGGG
               High quality sequence stop: 666.
FEATURES     Location/Qualifiers
               1..666
               /organism="Gossypium arboreum"
               /mol_type="mRNA"
               /strain="AKA"
               /cultivar="8400"
               /db_xref="taxon:29729"
               /clone="GA_Ba0012D16r"
               /tissue_type="Fibers isolated from bolls harvested 7-10
               dpa"
               /lab_host="E. coli"

```


Qy 411 PheGluAlaLysValAlaAspPheGlyLeuAlaLysIleAlaSerAspThrAsnThrHis 430
 Db 424 TTTGAAGCTAAGTGGCAGATTTTGGACTTCGAAAATTCACCTTCGTATACAAACACCCAT 483
 Qy 431 ValSerThrArgValMetGlyThrPheGlyTyrLeuAlaProGluTyrAlaAlaSerGly 450
 Db 484 GTTTCACACAGAGTAATGGGCACATTTGGGTACCTAGACACCTGAGTAGCTGCTTCGTGC 543
 Qy 451 LysLeuThrGluLysSerAspValPheSerPheGlyValValLeuLeuGluLeuIleThr 470
 Db 544 AAGCTCACTGAGAAATCAGATGCTCTCTCTTTGGAGTAAATGCTCTCTTGAGCTAATAACT 603
 Qy 471 GlyArgArgProValAspAlaAsnValTyrValAspAspSerLeuValAspTyrAla 490
 Db 604 GGGCGCGTCTCT-----GGTAATTCAGACAGCAGATGACAACTTTGGGTGACTGGGCA 657
 Qy 491 Arg-ProLeuLeuAsnArgAlaSerGluGlnGlyAsp 502
 Db 658 AGGCGCTTTGATGAATAAAGCATTTGAGGTGGTAAT 694

RESULT 15

BU099573

LOCUS

DEFINITION WHE3309_A09_A172S Chinese Spring wheat drought stressed root cDNA library Triticum aestivum cDNA clone WHE3309_A09_A17, mRNA sequence.

ACCESSION

BU099573

VERSION

EST.

KEYWORDS

Triticum aestivum (bread wheat)

SOURCE

Triticum aestivum

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

Pooideae; Triticeae; Triticum.

1 (bases 1 to 715)

Anderson, O.D., Chao, S., Close, T.J., Crossman, C., Fenton, R.D.,

Lazo, G.R., Nguyen H.T., Pham, J., Rausch, C.J., Turuspekov, Y.,

Wilson, C., Woo, J., and Zhang, D.

The structure and function of the expressed portion of the wheat

genomes - Chinese Spring drought stressed root cDNA library

Unpublished (2002)

Contact: Olin Anderson

US Department of Agriculture, Agriculture Research Service, Pacific

West Area, Western Regional Research Center

800 Buchanan Street, Albany, CA 94710, USA

Tel: 5105595773

Fax: 5105595818

Email: oanderson@w.usda.gov

Sequences have been trimmed to remove vector sequence and low

quality sequence with phred score less than 20

Seq primer: SK primer.

Location/Qualifiers

1..715

/organism="Triticum aestivum"

/mol_type="mRNA"

/cultivar="Chinese Spring"

/db_xref="taxon:4565"

/clone="WHE3309_A09_A17"

/tissue_type="root"

/dev_stage="Full tillering stage"

/lab_host="E. coli SOLR"

/clone_lib="Chinese Spring wheat drought stressed root

cDNA library"

/note="Vector: Lambda Uni-ZAP XR, excised phagemid;

Site 1: EcoRI; Site 2: XhoI; Plants were grown under

normal conditions, then drought stressed to 80%, 70% and

60% RWC at Texas Tech University (D. Zhang in Ht Nguyen

lab). Total RNA was prepared separately for roots

collected at the three different drought conditions. Equal

amounts of total RNA were pooled from all three samples,

poly(A) RNA were purified, one cDNA library was made, and

the cDNA clones were in vivo excised to give phluescript

SK(-) phagemids in the TJ Close lab at the University of

California, Riverside (Fenton, Turuspekov). Plasmid DNA preparations and DNA sequencing were performed in the OD Anderson lab (all other authors)."

ORIGIN

Alignment Scores: 1.04e-22 Length: 715
 Pred. No.: 983.00 Matches: 184
 Score: 89.70% Conservative: 25
 Percent Similarity: 78.97% Mismatches: 24
 Best Local Similarity: 28.47% Indels: 0
 Query Match: 13 Gaps: 0
 Db: 0

US-10-086-464-2 (1-647) x BU099573 (1-715)

Qy 276 SerGluAlaAsnLeuLeuGlyGlnGlyPheGlyTyrValHisLysGlyValLeuPro 295
 Db 11 TCCGATGCTAATCTTCTCGGGCAAGGTGGCTTTGGATTGTTTCAAAAGGAGTGTGCCA 70
 Qy 296 SerGlyLysGluValAlaValLysGlnLeuLysValGlySerGlyGlnGlyGluArgGlu 315
 Db 71 GATGGCAGAGAAGTTGCTGTGAAGCAATTAAGAGATGGAAGTGGGCGAGAGCGGTGAG 130
 Qy 316 PheGlnAlaGluValGluIleIleSerArgValHisArgHisLeuValSerLeuVal 335
 Db 131 TTCAGGCGAGAGGTGAGATTATCAGCCGAGTATCATATAACATCTCGTGACATTTGTT 190
 Qy 336 GlyTyrCysIleAlaGlyAlaLysArgLeuLeuValTyrGluPheValProAsnAsnAsn 355
 Db 191 GGTATTATGCAATTTCTGAAGACAAGAGGTGCTTGTCTATGAGTTTGTTCCTCAATAACAG 250
 Qy 356 LeuGluLeuHisLeuHisGlyGluGlyArgProThrMetGluTyrSerThrArgLeuLys 375
 Db 251 TTGATATTCCATATACATGGAAGCGGTGACCAACTATGGAAGTGGCTTCAAGACTACGT 310
 Qy 376 IleAlaLeuGlySerAlaLysGlyLeuSerTyrLeuHisGluAspCysAsnProLysIle 395
 Db 311 ATTGCTTTGGGTTCTCGAAGGGGATGGCGTATCTTCACGAAGACTGCGATCCAAAGATC 370
 Qy 396 IleHisArgAspIleLysAlaSerAsnIleLeuIleAspPheLysPheGluAlaLysVal 415
 Db 371 ATTATCGTGACATAAAGGCATCAATATTTCTTCTGGATTACAGATGGAAGCTAAGGTG 430
 Qy 416 AlaAspPheGlyLeuAlaLysIleAlaSerAspThrAsnThrHisValSerThrArgVal 435
 Db 431 GCAGATTTTGGACTTGCAAAAGTTAACTCTGATAATAACACTCATGTGTTCCACAGAGTA 490
 Qy 436 MetGlyThrPheGlyTyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLys 455
 Db 491 ATGGGCACATTTGGGTACCTTGCACAGAGTATGCTTCTTCTGCAAGCTAACTGAGAAA 550
 Qy 456 SerAspValPheSerPheGlyValValLeuLeuGluLeuIleThrGlyArgArgProVal 475
 Db 551 TCAGATGCTCTTTCTTTTGGAGTAATGCTTCTCGAGTTAATAACTGGGCGCGTCTCTGTA 610
 Qy 476 AspAlaAsnValTyrValAspAspSerLeuValAspTyrAlaArgProLeuLeuAsn 495
 Db 611 AGTTCAAAACAAGCGCATATGATGACAGCTTGGTTGCTGACTGGGCAAGCGCTTTGATGACA 670
 Qy 496 ArgAlaSerGluGlnGlyAspPheGluGlyLeuAlaAsp 508
 Db 671 CAAGCACTCGAGGATGATGAATACGATGCTTTAGTGAT 709

Search completed: May 13, 2004, 01:33:29

Job time : 3622 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus p2n model

Run on: May 12, 2004, 22:28:46 ; Search time 139 Seconds
(without alignments)
2583.117 Million cell updates/sec

Title: US-10-086-464-2
Perfect score: 3453
Sequence: 1 MSSAPSGTGTGPPSPNST.....REMEMGKIKRTGQYSGPSL 647

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlip
-Q=/cgn2_1/USPTO_spool_p/US10086464/runat_11052004_121815_7866/app_query.fasta_1.839
-DB=issued Patents NA -QMT=fastcap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10086464 -CGN 1 1.85 @runat_11052004_121815_7866 -NCPU=6 -ICPU=3
-NO MMAP -LARGESQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PTUS_COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	756	21.9	3239	4	US-09-228-986-9
2	744.5	21.6	1593	4	US-09-602-472A-1
3	740	21.4	2868	4	US-09-228-986-4
4	712	20.6	1488	4	US-09-579-182-4
5	703.5	20.4	2513	4	US-09-228-986-13
6	678	19.6	2432	4	US-09-228-986-7
7	667.5	19.3	1559	4	US-09-602-472A-3
8	604.5	17.5	4104	3	US-08-881-706-1
9	602	17.4	1224	4	US-09-579-182-3
10	600.5	17.4	3097	4	US-09-228-986-11
11	581.5	16.8	3590	1	US-08-587-889-1
12	581.5	16.8	3590	4	US-09-016-434-1093

13	581.5	16.8	3590	5	PCT-US96-09193-1	Sequence 1, Appli
14	578	16.7	2114	4	US-09-602-472A-5	Sequence 5, Appli
15	575	16.7	2749	1	US-08-265-628-1	Sequence 1, Appli
16	572	16.6	2571	1	US-07-717-331F-9	Sequence 9, Appli
17	572	16.6	2833	1	US-07-717-331F-1	Sequence 1, Appli
18	567	16.4	1554	2	US-08-587-680A-24	Sequence 24, Appli
19	566.5	16.4	2389	4	US-09-228-986-1	Sequence 1, Appli
20	562	16.3	2336	4	US-09-228-986-10	Sequence 10, Appli
21	559.5	16.2	2638	4	US-09-228-986-8	Sequence 8, Appli
22	546	15.8	2749	1	US-07-717-331F-4	Sequence 4, Appli
23	545	15.8	966	1	US-08-447-185-2	Sequence 2, Appli
24	543	15.8	2443	1	US-08-447-185-3	Sequence 3, Appli
25	543	15.7	2686	4	US-09-228-986-3	Sequence 3, Appli
26	531.5	15.4	5733	2	US-08-473-553A-1	Sequence 1, Appli
27	486	14.1	2943	4	US-09-503-922-2	Sequence 2, Appli
28	460.5	13.3	2649	4	US-09-228-986-12	Sequence 12, Appli
29	430	12.5	2568	4	US-09-228-986-2	Sequence 2, Appli
30	401	11.6	1926	4	US-09-249-585A-2	Sequence 2, Appli
31	401	11.6	1926	4	US-09-410-399-3	Sequence 3, Appli
32	401	11.6	2580	3	US-09-050-863-2	Sequence 2, Appli
33	401	11.6	2580	4	US-09-359-081-2	Sequence 2, Appli
34	401	11.6	5452	2	US-09-130-114-1	Sequence 1, Appli
35	401	11.6	8705	4	US-09-647-344A-14	Sequence 14, Appli
36	401	11.6	9600	3	US-08-910-647-1	Sequence 1, Appli
37	401	11.6	9600	4	US-09-620-925-1	Sequence 1, Appli
38	401	11.6	10596	1	US-07-884-811-15	Sequence 15, Appli
39	401	11.6	10596	1	US-07-885-971-15	Sequence 15, Appli
40	401	11.6	10596	1	US-08-087-783A-15	Sequence 15, Appli
41	401	11.6	10596	1	US-08-134-088B-15	Sequence 15, Appli
42	401	11.6	10596	2	US-08-194-087-15	Sequence 15, Appli
43	401	11.6	10596	5	PCT-US93-04648-15	Sequence 15, Appli
44	401	11.6	16080	4	US-09-724-566A-48	Sequence 48, Appli
45	395.5	11.5	4403765	3	US-09-103-840A-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1

US-09-228-986-9
; Sequence 9, Application US/09228986
; Patent No. 6359198
; GENERAL INFORMATION:
; APPLICANT: Nieuwenhuizen, Timothy
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signalling
; FILE REFERENCE: 11000/1020
; CURRENT APPLICATION NUMBER: US/09/228,986
; CURRENT FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 130
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 9
; LENGTH: 3239
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
; US-09-228-986-9

Alignment Scores:
Pred. No.: 1.88e-23
Score: 756.00
Percent Similarity: 51.93%
Best Local Similarity: 34.99%
Query Match: 21.89%
DB: 4
Matches: 3239
Conservative: 190
Mismatch: 92
Indels: 132
Gaps: 17

US-10-086-464-2 (1-647) x US-09-228-986-9 (1-3239)

Qy	89	ProProAlaProVal-ThrProThrArgAsnProProSerValProGlyProPr	108
Db	1759	CTCCCAACCATTCGCCCTTACTTATTCATCCCTCCCTATGCTTTT-----CC	1811
Qy	108	oSerAsnProSerArgGluGlySerProArgProProSerProSerProSe	128

```
Db 1812 AGACAAAT-----GGAGGAACC----- 1827
Qy 128 rProSerSerAspGlyLeuSerThrGlyValValValGlyLeuAlaLeuGlyValAl 148
Db 1828 -----GCCATAAGCAAAAGGTGTGATAGTTGGGATCGCTATTGGCGCACGGT 1874
Qy 148 aLeuLeuValValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 168
Db 1875 TCTGGTTCTTGGCCTTGTGTATTAGGGTTATATGCTATTTCGACAAAAGAACGGGCGGA 1934
Qy 168 uGluAspAlaTyrValProProPro-----ProProProGlyProLysAl 184
Db 1935 GAAA-----GCTCTCAGTTGACACACCTTCGCATCTCGGCACCCAGTGGAAGATAG 1991
Qy 184 aGlyGlyProTyrGlyGlyGlnGlnGlnIntrArgGlnGlnAsnAlaThrProProSe 204
Db 1992 CGGAGGAGCGCCACACACTGMAAGGACGACGATGG----- 2025
Qy 204 rAspHisValValThrSerLeuProProProProLysAlaProSerProProArgGlnPr 224
Db 2025 ----- 2025
Qy 224 oProProProProProProPheMetSerSerSerGlyGlySerAspTyrSerAspAr 244
Db 2025 ----- 2025
Qy 244 gProValLeuProProProSerProGlyLeuValLeuGlyPheSerLysSerThrPheTh 264
Db 2026 -----TTCTC 2030
Qy 264 rTyrGluGluLeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGl 284
Db 2031 CTATGATGAACCTTAAGAGGTGCACCAATAATTTCTCCGATAGCAATGAATTAGGCTTCGG 2090
Qy 284 yGlyPheGlyTyrValHisLysGlyValLeuProSerGlyLysGluValAlaValLysGl 304
Db 2091 AGGATACGGAAGGTGTACAGGGAGTCTCTCTGATGGTCATATATAGCAATCAAAAG 2150
Qy 304 nLeuLysValGlySerGlyGlnGlyGluArgGluPheGlnAlaGluValGluLeuLeuSe 324
Db 2151 AGCTCAGCAGCGGTTCGATCGAGGTGCAACCGAGTTCGAAGACAGAAATCGAGCTGCTTC 2210
Qy 324 rArgValHisHisArgHisLeuValSerLeuValGlyTyrCysLeuAlaGlyAlaLysAr 344
Db 2211 GCGGGTTTCATCACAAGAACTTGTGTGCGCTCATAGGATTCGTTTCGAGCAAGAGAGCA 2270
Qy 344 gLeuLeuValTyrGluPheValProAsnAsnAsnLeuGluLeuHisGlyGlyGluGl 364
Db 2271 GATGTTGGTCTATGATATATGCTTACGGGACGCTCAGGATAGCTTGACAGAAATC 2330
Qy 364 yArgProThrMetGluTrpSerThrArgLeuLysLeuAlaLeuGlySerAlaLysGlyLe 384
Db 2331 AGGCATTATCTTGTGGAAGAGAGGCTTCGTATAGCTCTAGGTTTCGGCTAGGAGCT 2390
Qy 384 uSerTyrLeuHisGluAspCysAsnProLysLeuLeuHisArgAspLeuLysAlaSerAs 404
Db 2391 AGCTTATCTGCACAACTCGCGAATCTCCAAATATTCACAGAGATGTCAGTCCACCAA 2450
Qy 404 nLeuLeuLeuAspPheLysPheGluAlaLysValAlaAspPheGlyLeuAlaLysLeuAl 424
Db 2451 TATCTTGTGGACCAACATCTGACGGCCAAAGTCGCGGATTCGTTTGTTCCAACTGGT 2510
Qy 424 aSerAspThrAsnThr-----HisValSerThrArgValMetGlyThrPheGlyTyrLeuAl 443
Db 2511 ATCGGACAGCGGGAAGGGGACGCTTCGACCAAGTGAAGGACGCTGGGCTATTGGA 2570
Qy 443 aProGlyTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPheSerPheGlyVa 463
Db 2571 TCCCGAATACATACATGATGATCAACAGCTGACAGAAAGAGCGATGTGTACAGCTTCGGGGT 2630
Qy 463 lValLeuLeuGluLeuLeuThrGlyArgArgProValAspAlaAsnValTyrValAs 483
Db 2631 GGTATGCTTGAGCTCATCTGCAAGCAACCAACGATTGAGNAGGCGCAAGTATGTGCTCGG 2690
```

```
Qy 483 pAspSerLeuValAspTrpAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPh 503
Db 2691 CGAG-----ATTTCGACCGCCATCGACAAGAACGACCCAG-----GACTA 2729
Qy 503 eGluGly-----LeuAlaAspAlaLysMetAsnAsn-----GlyTyrAspArgGluGl 519
Db 2730 CTACGGCGTGAAGAAATGATGGACCCGTCATCAGGAGCATGGGCTAC-----CTCGTCGG 2786
Qy 519 uMetAlaArgMetValAlaCysAlaAlaLysValArgHisSerAlaArgArgArgPr 539
Db 2787 GTTCAGCAGGTTCTTGATTTGGCGATGCGATGTGTCGAGGAGTCGCTCGGACCGCCC 2846
Qy 539 oArgMetSerGlnLeuValArgAlaLeuGluGlyAsnValSerLeuSerAspLeuAsnGl 559
Db 2847 CACAATGAGCGAGTGTGAAGCGATCGAGACCATGTTGCAG-----AACGA 2894
Qy 559 uGlyMetArgProGlyGlnSerAsnValTyrSerSerTyrGlyGlySerThrAspTyrAs 579
Db 2895 TGGGATACACACCAACTCGACGTCCTCGCATCGTCGCG-----GCGACGGACTTTGG 2945
Qy 579 pSerSerGln-----TyrAsnGluAspMetLysLysPheArgLysMe 593
Db 2946 GTCGACGAGGGCGCTCTCGGCATCCGTACACAGATCCCTTACCAG----- 2994
Qy 593 tAlaLeuGlyThrGlnGluTyrAsnAlaThrGlyGluTyrSerAsnProThrSerAspTy 613
Db 2995 -----AAGGAAGTTAGCTATAGCATTCCTTTGATTATAGT-----GGTGGATA 3038
Qy 613 rGlyLeu 615
Db 3039 TGGACTA 3045
```

RESULT 2

```
US-09-602-472A-1
; Sequence 1, Application US/09602472A
; Patent No. 6608240
; GENERAL INFORMATION:
; APPLICANT: Bidney, Dennis L.
; APPLICANT: Hu, Xu
; APPLICANT: Li, Guihua
; TITLE OF INVENTION: Sunflower Disease Resistance Genes
; FILE REFERENCE: 35718/200630
; CURRENT APPLICATION NUMBER: US/09/602,472A
; CURRENT FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/140,876
; PRIOR FILING DATE: 1999-06-24
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1593
; TYPE: DNA
; ORGANISM: Helianthus annuus
; FEATURE:
US-09-602-472A-1
```

```
Alignment Scores:
Pred. No.: 2,9e-23 Length: 1593
Score: 744.50 Matches: 172
Percent Similarity: 57.07% Conservative: 66
Best Local Similarity: 41.25% Mismatches: 116
Query Match: 21.56% Indels: 63
DB: 4 Gaps: 9
```

US-10-086-464-2 (1-647) x US-09-602-472A-1 (1-1593)

```
Qy 235 SerSerGlyGlySerAspTyrSerAspArgProValLeuProProSerProGlyLeu 254
Db 181 TCGTCGGGTTCGAAGACGCGGTTCGAGGGTCAGATT----- 216
Qy 255 ValLeuGlyPheSerLysSerThrPheThrTyrGluGluLeuAlaArgAlaThrAsnGly 274
Db 217 -----GCTGCTCATACATTCACTTCGCGGAGCTTGACAGTGCACAAACAAT 264
```



```
QY 287 GlyTyrValHisLysGlyValLeuProSerGlyLysGluValAlaValLysGlnLeuLys 306
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1188 GGAAGGTGTACAAAGGGCGCTTCGACATGGTCTCTGGTGGCTGTAAACGCTGAAG 1247
QY 307 ---ValGlySerGlyGlnGlyArgGluPheGlnAlaGluValGluLeuLeuLeuLeu 325
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1248 GAAGAGCCTACACCGGGTGGAGGTGCGAGTTTCAAACAGAGGTGGAGATGATGAAGCATG 1307
QY 326 ValHisHisArgHisLeuValSerLeuValGlyTyrCysIleAlaGlyAlaLysArgLeu 345
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1308 CGAGTACATAGGAACCTCTTCGACTAGTGGATTCTGCATGACACCACTGAACGGCTG 1367
QY 346 LeuValTyrGluPheValProAsnAsnLeuGlu-----LeuHisLeuHisGlyGlu 363
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1368 CTTGTTTATCCTACATGCCCAAGTGTGCTTCATCCCTACGAGAGAGGCGCA 1427
QY 364 GlyArgProThrMetGluTyrSerThrArgLeuLysIleAlaLeuGlySerAlaLysGly 383
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1428 AATGACCCCTTAGATTGGCAACTCGAAGCGCATAGCATTTGGGTCTGCAAGAGGT 1487
QY 384 LeuSerTyrLeuHisGlyAspCysAsnProLysIleIleHisArgAspIleLysAlaSer 403
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1488 CTCCTCTACTTGATGATCATTTGATCTTAAGATTATTCACCGGATGTCAAGGCTGCT 1547
QY 404 AsnIleLeuIleAspPheLysPheGluAlaLysValAlaAspPheGlyLeuAlaLysIle 423
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1548 AACACTTCTACTCGATGAAGATATGAGGACGTGTGGGGGATTTTGGCTTGGCAAACTT 1607
QY 424 AlaSerAspThrAsnThrHisValSerThrArgValMetGlyThrPheGlyTyrLeuAla 443
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1608 ATGGATTATAGGACACATGTTACGACGGCTGTTCTGTGAACCATTTGGCCACATGCA 1667
QY 444 ProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPheSerPheGlyVal 463
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1668 CTTGAGTACCTTTCTACTGGAAGCTCTCGGAAGACAGACGATTTATTTGGATATGGAATC 1727
QY 464 ValLeuLeuGluLeuIleThrGlyArgArgProValAspAlaAsnAsnValTyrValAsp 483
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1728 ATGTTGCTGGAACCTTATTACGGGACACACGGGCTTTGACCTTGCACGTTAGCAAAATGAT 1787
QY 484 AspSer-----LeuValAspTyrAlaArgProLeuLeuAsnArgAlaSerGluGlnGly 501
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1788 GATGATGTCATGTTCTTGACCTGGGTTAAAGCCCTACTA-----AAAGAGAGA 1835
QY 502 AspPheGlyGlyLeuAlaAspAlaLysMetAsnAsnGlyTyrAspArgGluGluMetAla 521
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1836 AGGCTTGATATGCTAGTTGATCCTCATCTTAAGAACAAATTTATTTGAAGCAGAGGTGGA 1895
QY 522 ArgMetValAlaCysAlaAlaCysValArgHisSerAlaArgHisArgArgProArgMet 541
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1896 CAACCTTATTCAAGTTGCTATTACTTTGTACACAAGGGTCCACCAATGGATAGACCAAGATG 1955
QY 542 SerGlnIleValArgAlaLeuGluGlyAsn-ValSerLeuSerAspLeuAsnGluGlyMet 561
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1956 TCTGAAGTGGTAAGATGTTGAAGGGGATGGCTTAGCTGAGAGATGGGAGATGGCA 2015
QY 561 TArgProGlyGlnSerAsnValTyrSerSerTyrGly 573
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
2016 AAGG-----TGGAGTCTGTACGGA 2034
```

RESULT 4

```
US-09-579-182-4
; Sequence 4, Application US/09579182
; Patent No. 6500628
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN KINASE AND
; FILE REFERENCE: PHOSPHATASE HOMOLOGUES AND USES THEREFOR
; CURRENT APPLICATION NUMBER: US/09/579,182
; CURRENT FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 7
```

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 4

; LENGTH: 1488

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-579-182-4

Alignment Scores:

Pred. NO.:	6.04e-22	Length:	1488
Score:	712.00	Matches:	178
Percent Similarity:	53.07%	Conservative:	73
Best Local Similarity:	37.63%	Mismatches:	160
Query Match:	20.62%	Indels:	62
DB:	4	Gaps:	14

US-10-086-464-2 (1-647) x US-09-579-182-4 (1-1488)

```
QY 205 AspHisValValThrSerLeuProPro-----ProProLysAlaProSer 219
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
22 GAACACAGATAAATCCCTTCGCTCCGTATATTCCTAAATCCGATAAATCAAAGTCT 81
QY 220 ProProArgGlnProProProProProProProProPheMetSerSerGlyGlySer 239
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
82 -----CGACGAGGTCT----- 93
QY 240 AspTyrSerAspArgProValLeuProProSerProGlyLeuValLeuGlyPheSer 259
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
94 -----GAACAGAGAAGGAGCTAACTGCTCCAAAGAGGGGCTACTGCGCATATTGCT 147
QY 260 LysSerThrPheThrTyrGluLeuAlaArgAlaThrAsnGlyPheSerGluAlaAsn 279
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
148 GCACAAACCTTACTTTCCGAGAGTTAGTCCGCCCTAAAAACTTTTCGACCGGAATGT 207
QY 280 LeuLeuGlyGlnGlyPheGlyTyrValHisLysGlyValLeu---ProSerGlyLys 298
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
208 CTTCTTGAGAAGGAGGTTTCGGACGTGTTTACAAAGGTCGCTAGAGACCACGAGACAG 267
QY 299 GluValAlaValLysGlnLeuLysValGlySerGlyGlnGlyGluArgGluPheGlnAla 318
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
268 ATAGTAGCTGTAAACAGCTTGATCGAAACCGTCTTCAAGGAAACAGAGAGTTTCTGTA 327
QY 319 GluValGluIleLeuSerArgValHisHisArgHisLeuValSerLeuValGlyTyrCys 338
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
328 GAGGTTCTTAGCTGAGCCTTCGCATCATCCCAATCTTGTGAATTGATTGTTATTGT 387
QY 339 IleAlaGlyAlaLysArgLeuLeuValTyrGluPheValProAsnAsnAsnLeuLeu 358
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
388 GCTGATGGGACACGAGCTCTTCTGTGTATGATATATGCCACTAGGATCATTCGAGGAT 447
QY 359 HisLeuHisGly-----GluGlyArgProThrMetGluTyrSerThrArgLeuLysIle 376
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
448 CATCTACACGATCTTCCACCTGATAAAGAGCCTCTAGACTGGAGTACTAGAAATGACAATA 507
QY 377 AlaLeuGlySerAlaLysGlyLeuSerTyrLeuHisGluAspCysAsnProLysIleIle 396
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
508 CGCGCAGGACGACGAAGGAGCTGGAGTATCTGCATGATAAAGCGAATCCGCTGTGATC 567
QY 397 HisArgAspIleLysAlaSerAsnIleLeuLeuAspPheLysPheGluAlaLysValAla 416
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
568 TACAGAGACCTGAAATCATCCCAACATCTTCTCGTGTATGGCTATCACCCAAAGTTATCT 627
QY 417 AspPheGlyLeuAlaLysIleAlaSer-----AspThrAsnThrHisValSerThrArgVal 435
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
628 GATTTGGGTTAGCTTAAGTTAGTCCCGTGGCGATAAACACATGTTGTCACCTCGTGTG 687
QY 436 MetGlyThrPheGlyTyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLys 455
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
688 ATGGGCATATGTTTATTGTGACCGGAAATATGCCATGACAGGGCACTCACATTGAAA 747
QY 456 SerAspValPheSerPheGlyValValLeuLeuGluLeuIleThrGlyArgProVal 475
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
748 TCCGATGTTTATAGCTTTGGGGTTGTGTTTCTCGAGCTCATACGGGTCGAAAAGCTATT 807
```

```
Qy 476 AspAlaAsnValTyrValAspSerSerLeuValAspTyrAlaAspProLeuLeuAsn 495
Dy 808 GATATGCTCGAGCACCCGGAGCACACCTTGTGCGATGGCTAGGCTGTTTC--- 864
Qy 496 ArgAlaSerGluGlnGlyAspPheGluGlyLeuAlaAspAlaLysMetAsnGlyTyr 515
Dy 865 -----AAAGATCGTAGAAGTTTCCGAAGATGGCGGATCCATCGCTCAAGGGCGGTAT 918
Qy 516 AspArgGluMetAlaArgMetValAlaCybAlaAlaCybValArgHisSerAla 535
Dy 919 CCAATGCTGCTATATCAAGACTTGCAGTGTGCAAGCAATGTGTTACAGAAACAGCA 978
Qy 536 ArgArgProArgMetSerGlnLeuValArgAlaLeuGluGlyAsnValSerLeuSer 555
Dy 979 GCGACAAGACCACTGATTGGGACGTGTGACAGCTCTTAACATATTAGCTTCGCAACG 1038
Qy 556 ---AspLeuAsn-----GluGlyMetArgProGlyGlnSerAsnValTyr--- 569
Dy 1039 TTTGACCAAAAGCCACCAAGCGGTCAAAACAGTAGAAGTGGGAGTGGGCCACCATTTATC 1098
Qy 570 -----SerSerTyrGlyGlySerThrAspTyrAsp----- 579
Dy 1099 AGAACAGGATGATCGGAGAGCTTGGGAGATGGGAGTAGCTGGATAGTCTTCGACAG 1158
Qy 580 -----SerSerGlnTyrAsnGluAspMetLysLysPheArgLysMet 593
Dy 1159 ACTCGAGTCGTTAGGCTACCGACCTACCAAGAACTCTCTGATTACAGAAGAGG 1218
Qy 594 AlaLeuGlyThrGlnGluTyrAsnAlaThrGlyGluTyrSerAsnProThrSer----- 611
Dy 1219 GATATG---GTGAGGGAAGTCAATGCAGGATCAGAAGGTGGGAGCGAGACAGGAGCGGG 1275
Qy 612 -----AspTyrGlyLeuTyrProSerGlySerSerSerGluGlyGlnThrArg 628
Dy 1276 TCAGTAGAAATGGGGATTA-----AGCGATTGGAAGGCGAAGAATCACAG 1323
Qy 629 GluMetGluMetGlyLysIleLysArgThrGlyGlnGly 641
Dy 1324 AGAGGAGCCCGCGGAGTGTGGGAGATCATCGAGAGGC 1362
```

RESULT 5

```
US-09-228-986-13
; Sequence 13, Application US/09228986
; Patent No. 6359198
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signalling
; FILE REFERENCE: 11000/1020
; CURRENT APPLICATION NUMBER: US/09/228,986
; CURRENT FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 130
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 13
; LENGTH: 2513
; TYPE: DNA
; ORGANISM: Pinus radiata
; US-09-228-986-13
```

```
Alignment Scores:
Pred. No.: 2,21e-21 Length: 2513
Score: 703.50 Matches: 178
Percent Similarity: 49.52% Conservative: 81
Best Local Similarity: 34.03% Mismatches: 179
Query Match: 20.37% Indels: 85
DB: 4 Gaps: 14
```

US-10-086-464-2 (1-647) x US-09-228-986-13 (1-2513)

```
Qy 84 ThrThrProGlySerProAlaProValThrProThrArgAsnProProSer 103
Dy 795 ACTGTCCTGCTCCTCTCTCCCAACCATATATCTCCACCT----- 833
```

```
Qy 104 ValProGlyProProSerAsnProSerArgGluGlyGlySerProArgProProSerSer 123
Dy 834 -----CCACCTTCTTAC-----AATCGGCCACCCCAAGTCC 863
Qy 124 ProSer-----ProProSerProSerSerAspGly-----Leu 134
Dy 864 TCAAAATGCTGGTCCCTTTCATCAGTGGGTCAAAAGGTCGGAACAGCAATAAGAAATCTCTG 923
Qy 135 SerThrGlyValValValGlyIleAlaIleGlyValAlaLeuLeuValIleValThr 154
Dy 924 AGTGGTGGTGCATAGTGGGTATAATATTTCAGCTTATTTCAGCTGTGTGTGCTGCTATA 983
Qy 155 LeuIleCysLeuLeuCysLysLysLysArgArgArgAspGluGluAspAlaTyrTyrVal 174
Dy 984 TTAGAGTTATTTTATGACAGTAACTCTCTTAGAAGAGAGCAGGAT----- 1031
Qy 175 ProProProProProGlyProLysAlaGlyGlyProTyrGlyGlnGlnGlnGln 194
Dy 1032 -----GAAGAAAACTA 1043
Qy 195 TrpArgGlnGlnAsnAlaThrPro-----ProSerAspHisValValThrSerLeuPro 212
Dy 1044 AGCAATCGTGTGTCTTTCACCCCTCTATCTCCCTCGATGCTGAATTATTGAAAGAGAGT 1103
Qy 213 ProProProLysAlaProSerProArgGlnProProProProProProProPhe 232
Dy 1104 CCAGAGCAAAAGTCAGCTATCACTCTTGAAATCGCTCTTAAGCTTCCCTCTCTGAA 1163
Qy 233 MetSerSerSerGly-----GlySerAspTyrSerAspArgProValLeu 247
Dy 1164 CGCAACAAGTCTACAGGGGACAAAGGCTTCGGAAGTATTTTTTCAAGTAAGAGACATAA 1223
Qy 248 ProProProSerProGlyLeuValLeuGlyPheSerLysSerThrPheThrTyrGluGlu 267
Dy 1224 AACCCA-----ATATCAACAACCTGAATATTCTATTTCAGAC 1259
Qy 268 LeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyPheGly 287
Dy 1260 CTGCAATATGCAACAAATAGTTTGTAGTCAAGATATCTTATTGCGGGGGTCTCTTGA 1319
Qy 288 TyrValHisLysGlyValLeuProSerGlyLysGluValAlaValLysGlnLeuLysVal 307
Dy 1320 CGAATCTACCGAGCAGAGTTCAGATGGAAGATTTTGGCAGTGAAGAAATTTGGACACT 1379
Qy 308 Gly-----SerGlyGlnGlyGluArgGluPheGlnAlaGluValGluIleSerArg 325
Dy 1380 TCTACGCTGTCTCCCTACAAAGCCCTGAAGACTTCTCGACCGCAGTATCTAATATATCGGC 1439
Qy 326 ValHisHisArgHisLeuValSerLeuValGlyTyrCysIleAlaGlyAlaLysArgLeu 345
Dy 1440 CTACATCATCTTAACATTACAGAACTAGTGGTTATTTCACAGAACATGACCATACCTT 1499
Qy 346 LeuValTyrGluPheValProAsnAsnAsnLeu-----GluLeuHisLeuHisGlyGlu 363
Dy 1500 CTTGTGTATGAATATTTTCGACAAATGGATCACTCTATCAGCTATTTCACATGCGAGATGAG 1559
Qy 364 GlyArgProThrMetGluTrpSerThrArgLeuLysIleAlaLeuGlySerAlaLysGly 383
Dy 1560 ACTACTAGAAATTTGCTCTGGAACATTCGTGTAAAGATTTGCGTGGGTTCAGCTCGAGTT 1619
Qy 384 LeuSerTyrLeuHisGluAspCysAsnProLysIleIleHisArgAspIleLysAlaSer 403
Dy 1620 TTAGAGTATTGTCATGAAAGTTTGTCTCTCATCTATTGTGCAATAAAAAATTCAGTCTCT 1679
Qy 404 AsnIleLeuIleAspPheLysPheGluAlaLysValAlaAspPheGlyLeuAlaLysIle 423
Dy 1680 AATATTTTGTGATGATGATTTCACCCCTCGTCTGCGCAGCTGTGGAAATTCGCG----- 1733
Qy 424 AlaSerAspThrAsnThrHisValSerThrArgValMetGlyThrPheGlyTyrLeuAla 443
Dy 1734 GCCCTCAATCAAATCTCGAGCGTCAGGTTTCAGGTGCTGGGTTCATTGGGTACAGTCT 1793
```

Qy	444	ProGluTyrAlaAlaSerGlyLeuThrGluLysSerAspValPheSerPheGlyVal	463
Db	1794	CCTGAATATGTCATGTCAGGAATCTATACATGAAGACGATGTTATAGTTTGGAGTG	1853
Qy	464	ValLeuLeuGluLeuIleThrGlyValArgProValAspAlaAenAenValTyrValAsp	483
Db	1854	GTAATGCTTGAGCTTTTGACAGGCCGGAGCCCTGGATAGTTCAGAAGACAGGTTCAGAA	1913
Qy	484	AspSerLeuValAspTyrAlaArgProLeuLeuAenArgAlaSerGluGlnGlyAspPhe	503
Db	1914	CAGTCATTAGTAAGATGGCCACCCCTCAGCTTCAT-----GATATT	1955
Qy	504	GluGlyLeuAla-----AspAlaIysMetAsnAsnGlyTyrAspArgGluGluMet	520
Db	1956	GATGCATTGGCAAGATGGTTGATGCCAGCTTTGAAGGGAAGTTATCTCGCAAAATCTCTC	2015
Qy	521	AlaArgMetValAlaCysAlaAlaCysValArgHisSerAlaAlaGArgAtgProArg	540
Db	2016	TCGGCGCTTGCTGATATTATTGGCCCTCTGCATTCCAGCCTGAACCAGAAATCCCGTCCCTCA	2075
Qy	541	MetSerGlnIleValArgAlaLeuGluGlyAsnValSerLeuSerAspLeuAsnGluGly	560
Db	2076	ATGTCGAAGTGGTGCAAGCATTTGGTTCTGATGATGCAACGGGTCAGCTCAATTAAGAGG	2135
Qy	561	MetArgProGlyGlnSerAsnValTyrSerSerTyrGlyGlySerThrAspTyrAspSer	580
Db	2136	ATGCACGAGNATGAA-----ACTGCAGACACGACATCTCT	2168
Qy	581	SerGlnTyr	583
Db	2169	GCAGATTAT	2177

RESULT 6

```

RES001 8
US-09-228-986-7
; Sequence 7, Application US/09228986
; Patent No. 6359198
; GENERAL INFORMATION:
; APPLICANT: Strábala, Timothy
; APPLICANT: Nieuwenhuizen, Niels
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and their Use in the Modification of Plant Cell Signalling
; FILE REFERENCE: 11000/1020
; CURRENT APPLICATION NUMBER: US/09/228,986
; CURRENT FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 130
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 2432
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-228-986-7

```

US-10-086-464-2 (1-647) x US-09-228-986-7 (1-2432)

Qy	254	LeuValLeuGlyPheSerLysSerThrPheThrTyrGluGluLeuAlaArgAlaThrAsn	273
		:::	
Db	1046	ATTTGTC TAGGCATTTCAAG--AGATTCTCAATCAAAGAGTTACGTGTGTTCCACTAAT	1102
		:::	
Qy	274	GlyPheSerGluAlaAsnLeuLeuGlyGlnGlyGlyPheGlyTyrValHisLysGlyVal	293
Db	1103	AATTTTAGTACCAAGAAATATTTTAGGAGTAGGAGGATATGCAATGCTCATAAAGGATTC	1162
Qy	294	LeuProSerGlyLysGluValAlaValLysGlnLeuLysValGlySer---GlyGlnGly	312
Db	1163	CTCAAGATGTCACCTATAGTAGCAATATAAAGGTTGAAGATGCTAATGTGGGAGGAGGA	1222

RESULT. T 7

```

RESULT 7
US-09-602-472A-3
; Sequence 3, Application US/09602472A
; Patent No. 6508240
; GENERAL INFORMATION:
; APPLICANT: Bidney, Dennis L.
; APPLICANT: Hu, Xu
; APPLICANT: Li, Guihua
; TITLE OF INVENTION: Sunflower Disease Resistance Genes
; FILE REFERENCE: 35718/200630
; CURRENT APPLICATION NUMBER: US/09/602,472A
; CURRENT FILING DATE: 2000-06-23

```



```

QY 158 Leu-----LeuCysLysLysLysLysArgArgArgAspGlu 168
DB 2515 ATATTGGGCGTATCTTTGGTAGAGAGATGAGAAAGAGAGAGAGAGAGCG 2574
QY 169 GluAspAlaTyrTyrValProProProProProGlyProLysAlaGlyGlyProTyr 188
DB 2575 GAGTTGGAGATGTATGCG-----GAAGGACATGGAAACTCTGGCGATAGA 2619
QY 189 GlyGlyGlnGlnGlnTrpArgGlnGlnAsnAlaThrProProSerAspHisValVal 208
DB 2620 ACTGCTAACACACCAATTTGGAAGCTGACTGGTGTGAAGAAGCCTTGAGTATCAATCTT 2679
QY 209 ThrSerLeuProProProLysAlaProSerProProArgGlnProProProPro 228
DB 2680 GCTGCTTTTCGAAGCCATTGCG----- 2703
QY 229 ProProProPheMetSerSerSerGlySerAspTyrSerAspArgProValLeuPro 248
DB 2703 ----- 2703
QY 249 ProProSerProGlyLeuValLeuGlyPheSerLysSerThrPheThrTyrGluGluLeu 268
DB 2704 -----AAGCTCACGTTTGGCGATCTT 2724
QY 269 AlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyPheGlyTyr 288
DB 2725 CTTGAGGCTACCAATGGTTTCATAATGATAGTCTGTGTTCTGTTGGTGGAGAT 2784
QY 289 ValHisLysGlyValLeuProSerGlyLysGluValAlaValLysGlnLeuLysValGly 308
DB 2785 GTTTACAAGCGATTTTGAAGATGGAAGCGGTGGCTATCAAGAAACTGATCATGTT 2844
QY 309 SerGlyGlnGlyGluArgGluPheGlnAlaGluValGluLeuLeuSerArgValHisHis 328
DB 2845 AGCGGTCAAGGTGATAGAGATTTCATGCGGAGATGGAACCATTTGGGAAGATCAACAT 2904
QY 329 ArgHisLeuValSerLeuValGlyTyrCysLeuAlaGlyAlaLysArgLeuValTyr 348
DB 2905 CGAAATCTTGTCCTCTCTTGGTTATTGCAAGTTGGAGACGAGCGGCTTCTGTTAAT 2964
QY 349 GluPheValProAsnAsnLeuGlu-----LeuHisLeuHisGlyGluGlyArgPro 366
DB 2965 GAGGTATGAAGTATGGAAGTTTAGAAGATGTTTTCGAAGACCCCAAGAAAGGTGGGGTG 3024
QY 367 ThrMetGluTrpSerThrArgLeuLysIleAlaLeuGlySerAlaLysGlyLysSerTyr 386
DB 3025 AAACCTAAATTCCTCACACGCGGAAGATTGCGATAGGATCAGCTAGAGGCTTGCTTTC 3084
QY 387 LeuHisGluAspCysAsnProLysIleIleHisArgAspIleLysAlaSerAsnLeuLeu 406
DB 3085 CTTACCAACAACGTGCGATATCATCCACAGAGACATGAATCCAGTAATGTGTGTG 3144
QY 407 IleAspPheLysPheGluAlaLysValAlaAspPheGlyLeuAlaLysIleAlaSerAsp 426
DB 3145 CTTGATGAGAATTGGAAGCTCGGGTTTCAGATTGTCATGTCGCGAGCTGATAGTGGC 3204
QY 427 ThrAsnThrHisValSerThrArgValMet---GlyThrPheGlyTyrLeuAlaProGlu 445
DB 3205 ATGGATACGCAATTAAAGCGTCAGTACATTAGCTGGTACACCGGGTTACGTTCTCCAGAG 3264
QY 446 TyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPheSerPheGlyValValLeu 465
DB 3265 TATTACCAAAAGTTTCAGGTGTTTCAACAAAAGGAGACGTTTATAGTTACGGTGTGTCCTTA 3324
QY 466 LeuGluLeuIleThrGlyArgArgProValAspAlaAsnValTyrValAspAspSer 485
DB 3325 CTCGAGCTACTACCGGGTAACCGGCCAACCGGATTCACCGGAT---TTGGAGATAAACAC 3381
QY 486 LeuValAspTrpAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGluGly 505
DB 3382 CTTGTTGGTGGTGAACAG----- 3402

```

```

QY 506 LeuAlaAspAlaLysMetAsnAsnGlyTyrAspArgGlu----- 518
DB 3403 CACGCAAAACTGCGGATTAAGCATGTGTTGACCCGAGGCTTATGAAGGAAGATCCAGCA 3462
QY 519 -----GluMetAlaArgMetValAlaCysAlaAlaAlaCysValArgHisSerAla 535
DB 3463 TTAGAGATCGAACCTTTTACAACATTTAAAGTTGCGGTGCGTGTGGATGATCGGGCT 3522
QY 536 ArgArgArgProArgMetSerGlnIleValArgAlaLeuGluGlyAsnValSerLeu--- 554
DB 3523 TGGAGACGACCGACAATGGTACAGTCAAGTCATGGCCATGTTTAA-GGAGATACAAGCGGGTC 3581
QY 555 -----SerAspLeuAsnGluGlyMetArgProGlyGlnSerAsn 567
DB 3582 AGGATAGATTACAGTCAACGATCAGATCAATCAGAGGATGGAGG----- 3626
QY 568 ValTyrSerSerTyrGlyGlySerThrAspTyrAspSerSerGlnTyrAsnGluAsp 586
DB 3627 -----GTTTCAGTACATAGAGAT 3644

RESULT 9
US-09-579-182-3
; Sequence 3, Application US/09579182
; Patent No. 650628
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN KINASE AND
; FILE OF INVENTION: PHOSPHATASE HOMOLOGUES AND USES THEREFOR
; FILE REFERENCE: MNI-161
; CURRENT APPLICATION NUMBER: US/09/579,182
; CURRENT FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1224
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-579-182-3

Alignment Scores:
Pred. No.: 1,81e-17 Length: 1224
Score: 602.00 Matches: 153
Percent Similarity: 48.41% Conservative: 60
Best Local Similarity: 34.77% Mismatches: 157
Query Match: 17.43% Indels: 70
DB: 4 Gaps: 9

US-10-086-464-2 (1-647) x US-09-579-182-3 (1-1224)
QY 128 SerProSerSerAspGlyLeuSerThrGlyValValValGlyIleAlaIleGlyGlyVal 147
DB 28 TCTCGGAACACTACAATGGTGTGGAGCTATGGAGATATAGTATGTTCTCTCCGCGGATA 87
QY 148 AlaLeuLeuValIleValThrLeuIleCysLeuLeuCysLysLysLysArgArgArgAsp 167
DB 88 TTGCTGTAGTTTTAGCTATATCGCTATATGCTTACTTTCAGAGAAAACCTCTAGATCT 147
QY 168 GluGluAsp-----AlaTyrTyrValProProProProProGlyProLys 183
DB 148 TCTTCTAATCTAATCTTGTAGTGGCGAGATTCCTCTCTAGTGTTCCTGAAGAGATATAA 207
QY 184 -----AlaGlyGlyProTyrGlyGlnGlnGlnGlnTrpArgGln 197
DB 208 GAGATTAGTTCGACGAGGTTTCTTCAAGCAATGCTGGG----- 246
QY 198 GlnAsnAlaThrProProSerAspHisValValThrSerLeuProProProPro 215
DB 247 ---AATGGATACCCCTCTATTAGTGAGAAATTTGGCGATAAAGAACCCGAAAAAGGGATA 303
QY 216 LysAlaProSerProProArgGlnProProProProProProProPheMetSerSer 235
DB 304 AAACGAGAGTCAGAAATGGCGATAGTAGCCGGTCAGGCTCGTTTAATCACTTCGAGAAA 363

```

```
Qy 236 SerGlyGlySerAspTyrSerAspArgProValLeuProProSerPro-----Gly 253
Db 364 AAAGACGGATCGAGGTATCTTCTGCTAATCCTTTGACAGCTCCATCTCTCTTGTCTGTT 423
Qy 254 LeuVal-----LeuGlyPheSerLysSerThrPheThrTyrGluGluLeuAla 269
Db 424 CTTCTCTGAGTTTCTCACCTTGGATGGGA---CATTTGGTTCACTCTTAGAGATCTTCAG 480
Qy 270 ArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyGlyPheGlyTyrVal 289
Db 481 ATGCTACTAATCAGTTTCAAGGATATATCATCGGTGATGTTGGATATGGAGTTGTT 540
Qy 290 HisLysGlyValLeuProSerGlyLysGluValAlaValLysGlnLeuLysValGlySer 309
Db 541 TACCGCGTAACCTTGTAAATGGTACTCTCTGCTGTTAAAGATTGCTCAACAATTTA 600
Qy 310 GlyGlnGlyGluArgGluPheGlnAlaGluValGluLeuLeuSerArgValHisHisArg 329
Db 601 GGACAAGCTGATAAAGACTTCAGAGTTTGAAGTTGAAGCTATAGGTTCAGTTCGACATAAA 660
Qy 330 HisLeuValSerLeuValGlyTyrCysIleAlaGlyAlaLysArgLeuLeuValTyrGlu 349
Db 661 AACTTGGTCCCTTCTCGGATATATGATGGAAGAACGACAGG-----705
Qy 350 PheValProAsnAsnLeuGluLeuHisGlyGlyGluGlyArgProThrMetGlu 369
Db 705 -----705
Qy 370 TrpSerThrArgLeuLysIleAlaLeuGlySerAlaLysGlyLeuSerTyrLeuHisGlu 389
Db 706 -----CTCGCGTACTCTTCAGGAG 723
Qy 390 AspCysAsnProLysIleIleHisArgAspIleLysAlaSerAsnIleLeuIleAspPhe 409
Db 724 GCGATTGAGCAAAAGTGTGTCAGACATTAAGTCTAGTAACATCTCATTTGATGATGAC 783
Qy 410 LysPheGluAlaLysValAlaAspPheGlyLeuAlaLysIleAlaSerAspThrAsnThr 429
Db 784 AAATTCATTAATTTCTGACTTTTGACTTTGGACTTGTCTAAACTACTTTGGTGTGTAAGAGT 843
Qy 430 HisValSerThrArgValMetGlyThrPheGlyTyrLeuAlaProGluTyrAlaAlaSer 449
Db 844 TTTATAACTACTAGATTATGGTACCTTCGGTTACGTAGCTCCAGAGATATCGCAATTCC 903
Qy 450 GlyLysLeuThrGluLysSerAspValPheSerPheGlyValValLeuLeuGluLeuIle 469
Db 904 GGTCTCTGAATGAGAAAGCGATGCTACAGCTTCGGGGTTGTACTCTTGGAAAGCTATA 963
Qy 470 ThrGlyArgArgProValAspAlaAsnAsnValTyrValAspAspSerLeuValAspTrp 489
Db 964 ACTGATAGATATCGGTAGACTATGCTCGTCCACCACCGAGGTACATTTGGTGGAGTGG 1023
Qy 490 AlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGlyGlyLeuAlaAspAla 509
Db 1024 CTGAAGATGATGGTC-----CAACAAAGACGATCAGAAGAGTGGTTGATCCA 1071
Qy 510 LysMetAsnAsnGlyTyrAspArgGluGluMetAlaArgMetValAlaCysAlaAlaAla 529
Db 1072 AACCTTGAACAAACCACTTACAAAGTCTTTGAAAAGAACACACTATTGACTGCTTTGAGA 1131
Qy 530 CysValArgHisSerAlaArgArgProArgMetSerGlnIleValArgAlaLeuGlu 549
Db 1132 TGTGTTGATCCAATGCTGAGAAAAGACCGAGGATGAGCCAGTTGACGATGCTTTGAA 1191
```

RESULT 10

```
US-09-228-986-11
; Sequence 11, Application US/09228986
; Patent No. 6359198
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Niels
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signalling
```

```
; FILE REFERENCE: 11000/1020
; CURRENT APPLICATION NUMBER: US/09/228,986
; CURRENT FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 130
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 3097
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-09-228-986-11

Alignment Scores:
Pred. No.: 4,95e-17 Length: 3097
Score: 600.50 Matches: 173
Percent Similarity: 46.08% Conservative: 80
Best Local Similarity: 31.51% Mismatches: 172
Query Match: 17.39% Indels: 126
DB: 4 Gaps: 13
```

```
US-10-086-464-2 (1-647) x US-09-228-986-11 (1-3097)

Qy 66 ProSerThrProGlySerProProPro-----LeuPro 76
Db 1523 CTTGCAAGTTCTGGACCTCCAAACACAGCTTAGACGGAGTTTCCCGACGGCTTGGG 1582
Qy 77 GlnProSerProAlaProThrThrProGlySerProAlaProValThrPro--- 95
Db 1583 AGAGCTGAAGACCTTCACCTACTGGACTTGGAGATAAAGCTGTACAG-GTACCTTAC 1641
Qy 96 -----ProThrArgAsnProProProSerValProGlyProPro 108
Db 1642 CAGATTCCTTGAACAGACAGAGCTTGGAGGTCAGAACCTCAGGAACCTTGTGCTTCT 1701
Qy 109 SerAsnProSerArgGluGlySerProArgProProSerProSerPro----- 126
Db 1702 TCTCCACAACCGCGTGC-GGTGATCATCTCTTCAATTGAGGCCACCCAAAGTT 1760
Qy 127 -----ProSerProSerSerAspGlyLeuSerThrGlyValValValGlyIleAla 143
Db 1761 ACATAGTTCCCGAGAGAAACAAAGGGGACATAATCGTTAGCCATATACTCGGAGCA 1820
Qy 144 IleGlyGlyValAlaLeuLeuValIleValThrLeuIleCysLeuLeuCysLysLys 163
Db 1821 GTCGAGGAGTATCAGCTAGCTATTTTACTATCCGCTCTTCGTATTTCATGACAGAAG 1880
Qy 164 ArgArgAspGluGluAspAlaTyrTyrValProProProProProGlyProLys 183
Db 1881 AGAGGAAGAACTGAA-----1895
Qy 184 AlaGlyGlyProTyrGlyGlyGlnGlnGlnTrpArgGlnAsnAlaThrProPro 203
Db 1895 -----1895
Qy 204 SerAspHisValValThrSerLeuProProProLysAlaProSerProProArgGln 223
Db 1895 -----1895
Qy 224 ProProProProProProPheMetSerSerSerGlySerAspTyrSerAsp 243
Db 1896 -----ATGTCATATACGAA 1910
Qy 244 ArgProValLeuProProProProGlyLeuValLeuGlyPheSerLysSerThrPhe 263
Db 1911 AGGCAGTCCAGACGCTGAGAAACTGGAAT-----GCAGCTAAGATTTT 1955
Qy 264 ThrTyrGluGluLeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGln 283
Db 1956 TCCTACAAAGAGATCAAAACAGCTACAAACAACTTTAAAGAA-----GTCAATGGTCA 2009
Qy 284 GlyGlyPheGlyTyrValHisLysGlyValLeuProSerGlyLysGluValAlaValLys 303
Db 2010 GGAAGTTTGGATCCGCTGATCGTGGGAACCTTCCAGTTGGAACAACTAGTTGCTGTGAA 2069
```

```
Qy 304 GlnLeuLysValGlySerGlyGlnGlyGluArgGluPheGlnAlaGluValGluLeuIle 323
Db 2070 GTGCGGTTTGATAAAACCACTTGGTGCAGATTCTTTTCATAAATGAGGTTGCTCTCTTA 2129
Qy 324 SerArgValHisHisArgHisLeuValSerLeuValGlyTyrCysIleAlaGlyAlaLys 343
Db 2130 TCACAACTGCCCATCAGAACCTTGTCAGTCTGGAGGATTTTGTATGAGTCGACGGT 2189
Qy 344 ArgLeuLeuValTyrGluPheValProAsnAsnLeuGluLeuHisLeuHisGlyGlu 363
Db 2190 CAGATTTTGTCTATGAATATCTACCGGTTGGATCACTGGCTGATCAACTGTATGGTCCA 2249
Qy 364 Gly-----ArgProThrMetGluTyrSerThrArgLeuLysIleAlaLeuGlySerAla 381
Db 2250 AACAGTAGGAATCTCTCACTAGCTGGGTCGTAGACTCAAGATTTGTTAGTGTGCA 2309
Qy 382 LysGlyLeuSerTyrLeuHisGluAspCysAsnProLysIleIleHisArgAspIleLys 401
Db 2310 AAAGACTGGACTATCTACATTAATGGAAGCAATCTCGAATCATACACCGAGACATCAAG 2369
Qy 402 AlaSerAsnIleLeuIleAspPheLysPheGluAlaLysValAlaAspPheGlyLeuAla 421
Db 2370 TGCAGTAATATCTATTGACAGAGATGAATGCAAGACTTTGCGACTTTGGGCTCTCT 2429
Qy 422 Lys-----IleAlaSerAspThrAsnThrHisValSerThrArgValMetGlyThrPhe 439
Db 2430 AAGCAAAATGATCCAGCCAGAGCGCA---ACTCACGTGACCACTGTCGTCAGGGGACAGCT 2486
Qy 440 GlyTyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPhe 459
Db 2487 GGTACCTCGACCTGATATATCTACCCACCAACTTACAGAGAAAGCGAGCTCTAT 2546
Qy 460 SerPheGlyValValLeuLeuGluLeuIleThrGlyArgArgProValAspAlaAsnAsn 479
Db 2547 AGCTTTGAGTTGTCTTTTGGAGCTCATCTGTGACGAGAGCCGTTAAATCAITTCAGGA 2606
Qy 480 ValTyrValAspAspSerLeuValAspTrpAlaArgProLeuLeuAsnArgAlaSerGlu 499
Db 2607 ACTCCAGATTCTCTCAATTTGGTTTATGGCAAGCCCTACTTG-----CAG 2654
Qy 500 GlnGlyAspPheGluGlyLeuAlaAlaLysMetAsnAsnGlyTyrAspArgGluGlu 519
Db 2655 CGAGGTGCATTGAG---ATAGTGATGAGAGTTTAGGGGAAGTTTCGATGTGGAAGC 2711
Qy 520 MetAlaArgMetValAlaCysAlaAlaCysValArgHisSerAlaArgArgPro 539
Db 2712 ATGAGAAAGTGGCAAAATCGCTGTGAGTCTGTAGAGGGATGCATCACTAAGGCCA 2771
Qy 540 ArgMetSerGlnIleValArgAlaLeuGluGly-----AsnValSerLeuSerAspLeu 557
Db 2772 ACCATTGCACAGATACTCTGTGCTCAAGAGGCTTACAGCAATTCAGCTCTCT----- 2825
Qy 558 AsnGluGlyMetArgProGlyGlnSerAsnValTyrSerSerTyrGlyLysThrAsp 577
Db 2826 -----TATCTTGAGCCTCTGGACATGTGAC 2852
Qy 578 TyrAspSerSerGlnTyrAsnGluAsp 586
Db 2853 TGAATCTACTCGTATTACCAACCAAGAC 2879
```

RESULT 11

US-08-587-889-1
; Sequence 1, Application US/08587889
; Patent No. 5654397

GENERAL INFORMATION:

; APPLICANT: CAO, Zhaodan
; APPLICANT: CROSTON, Glenn E.
; APPLICANT: GOEDEL, David V.
; TITLE OF INVENTION: INTERLEUKIN-1 RECEPTOR-ASSOCIATED
; TITLE OF INVENTION: PROTEIN KINASE AND BINDING ASSAY
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT

; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/587,889
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman Ph.D., Richard Aron
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: A-60916
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 494-8700
; TELEFAX: (415) 494-8771
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3590 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-587-889-1

Alignment Scores:
Pred. No.: 3,48e-16 Length: 3590
Score: 581.50 Matches: 208
Percent Similarity: 38.68% Conservative: 86
Best Local Similarity: 27.37% Mismatches: 239
Query Match: 16.84% Indels: 229
DB: 1 Gaps: 28

US-10-086-464-2 (1-647) x US-08-587-889-1 (1-3590)

```
Qy 1 MetSerAlaProSerPro-GlyThrGlySerProSerProSerProSerAsnSerTh 20
Db 80 ATGCCGGGGGCGCGCCCGGGGAGCCCGCAGACCCCGCGGCGCCAGCATTTCTGTAC 139
Qy 20 rThrThrThrProProAlaSerAla-----ProProPr 32
Db 140 GAGGTGCCGCGCTGGGTGTCATGTGCGGCTTCTACAAAGTGATGGACGCCCTGGAGCCGCC 199
Qy 32 oThrThrProSerSerProPro----- 39
Db 200 GACTGTGTGCCAGTTTCGCGCCCTGATCGTGCAGCACCGACCGAGCTGCGGCTGTGCGAG 259
Qy 40 ----ProProSerThrIleProThrSerProProProSerSerArgSerThrProSerAl 58
Db 260 CGCTCCGGGAGCCAGCCAGCGCGCTCTGTGGCCCTGGATCAACCGCAACGCCCGGTGTG 319
Qy 58 aProProProSerProPro-----ThrProSerTh 68
Db 320 GCCGACCTCGTGACATCTCTCAGCACCTGCGAGCTGCTCGTGCAGCGGACATCATCACA 379
Qy 68 rProGlySerProProProLeuProGlnProSerProProAlaProThrThrProGlySe 88
Db 380 GCCTGGGACCCCTCCCGCCCGCTTCCTGTCGCCAGCACCACTGCCCCGA-----GG 430
Qy 88 rProProAlaProValThrProThrArgAsnProProProSerValProGlyProPr 108
Db 431 CCCAGAGCATCTCTGCACCGCGG-----AGCGCGAGGCTGGAGCCCCCGGA----- 479
Qy 108 oSerAsnProSerArgGluGlySerProArgProProSerSerProSerProProSe 128
Db 480 -----AGTTGCCATCTCTCAGCTCCACCTTCTCTC-TCCCCAGCTTTT 519
```

Qy 128 rProSerSerAspGlyLeuSerThrGlyValValValValGlyLeuAlaLeuGlyValAl 148
Db 520 TCCAGGCTCCAGACCCATTCA----- 541
Qy 148 aLeuLeuValLeuValThrLeuLeuLeuCysLeuLeuLeuValValValValValVal 168
Db 541 ----- 541
Qy 168 uGluAspAlaTyrTyrValProProProProProGlyProLysAlaGlyGlyProTy 188
Db 542 -----GGGCTGAGCTGGC----- 556
Qy 188 rGlyGlyGlnGlnGlnTrpArgGlnGlnAsnAlaThrProProSerAspHisValVa 208
Db 556 ----- 556
Qy 208 lThrSerLeuProProProProProProProSerProProArgGlnProProProProPr 228
Db 557 -----CTGGTTCCAAAGCCCTGCTTCCCTGTGGCTCCACCGCC 594
Qy 228 oProProProPheMetSerSerGlySerAspTyrSerAspTyrProValLeuPr 248
Db 595 ATCTCCAGCCCTTCTTACCAAGCCAGCC-----CC 627
Qy 248 oProProSerProGlyLeuValLeuGlyPheSerLysSerThrPheThrTyrGlu----- 266
Db 628 AGAGAGCTCAGTGTCCCTCTGCAGGAGCCGCCCTCTCCGTTTGTGCTGGCCCTCTG 687
Qy 267 -GluLeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyGlyph 286
Db 688 TGAGATTTCCTGGGGGACCCCAACTTCTCGAGAGAGCTCAAGATCGGGAGGGTGGCTT 747
Qy 286 eGlyTyrValHisLysGlyValLeuProSerGlyLysGluValAlaValLysGlnLeuL 306
Db 748 TGGTGTGCTGTACCGCGCGTGATG-----AGGAACACCGTGTATGCTGTGAAGAGCTGAA 804
Qy 306 sValGlySerGly-----GlnGlyGluArgGluPheGlnAlaGluValGluI 322
Db 805 GGAGAACGCTGACCTGGAGTGCAGTGCAGTGAAGCAGAGCTTCTGACCGAGGTGGAGCA 864
Qy 322 eLysSerArgValHisArgHisLeuValSerLeuValGlyTyrCysLeuAlaGlyAl 342
Db 865 GCTGTCCAGGTTCGTACCCCAACATTTGTGGACTTTGCTGGCTACTGTCTCAGAACGG 924
Qy 342 aLysArgLeuLeuValTyrGluPheValProAsnAsnLeuGlu-----LeuHisLe 360
Db 925 CTCTACTGCTGGTGTACGGCTTCTCTGCCCAACAGCGCTCCCTGGAGACCGCTCTCCACTG 984
Qy 360 uHisGlyGlyGlyArgProThrMetGluTrpSerThrArgLeuLysLeuAlaLeuGlySe 380
Db 985 CCAGACCCAGCGCTGCCACCTCTCTCTGSCCTCAGCGACTGGACATCTCTCTGGGTAC 1044
Qy 380 rAlaLysGlyLeuSerTyrLeuHisGluAspCysAsnProLysLeuLeuHisArgAspI 400
Db 1045 AGCCCGGGCAATTTCAGTTTTCATCATCAGGAC---AGCCCCAGCGCTCATCTCATGAGACAT 1101
Qy 400 eLysAlaSerAsnLeuLeuLeuAspPheLysPheGluAlaLysValAlaAspPheGlyLe 420
Db 1102 CAAGATTCCAACTGCTTCTGGATGAGAGCTGACACCCCAAGCTGGGAGACTTTGGCCT 1161
Qy 420 uAlaLysLeu-----AlaSerAspThrAsnThrHisValSerThrArg-- 434
Db 1162 GGCCCGGTTACGCGCTTTGCGGGTCCAGCCCGCCAGCAGCAGTGTGGCGCCGAC 1221
Qy 435 -----ValMetGlyThrPheGlyTyrLeuAlaProGluTyrAlaAlaSerGlyLysLe 452
Db 1222 ACAGACAGTGGGGGACCCCTCTGCTGCTACCTGCCGAGGAGTACATCAAGACGGGAGGCT 1281
Qy 452 uThrGluLysSerAspValPheSerPheGlyValValLeuLeuLeuLeuLeuThrGlyAr 472
Db 1282 GGCTGTGACACGACACCTTTCAGCTTTGGGGTGGTGTGTGTAGAGACCTTGGCTGTCA 1341
Qy 472 gArgProVal-----AspAlaAsnAsnValTyrValAspAspSerLeuValAspTr 489

Db 1342 GAGGCTGTGAAGACGACCGTGTCCAGGACCAAGTATCTGAAAGAC---CTGTGGAA-- 1396
Qy 489 pAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGluGlyLeuAlaAspAl 509
Db 1397 -----GAGGAGGCTGAGGAGGCTGGAGTGGCTTTTGAGAGCACCCAGAG 1440
Qy 509 aLysMetAsnAsnGlyTyrAspArgGluGluMetAlaArgMetValAla----- 525
Db 1441 CACACTGCAAGCAGGTCTGCTGCAGATGCTGGGCTGCTCCCATCGCCATGCAGATCTA 1500
Qy 525 ----- 525
Db 1501 CAAGAAGCACCTGGACCCCGCGCCCTGCTGCCACCTCAGCTGGCGCTGGCGCTGGG 1560
Qy 526 ----CysAlaAlaAlaCysValArgHisSerAlaArgArgProArgMetSerGlnI 544
Db 1561 CCAGCTGCTGCTGCTGCTGCCTGCACCCCGGGGCAAAAGAGGCTCTCTATGACCCAGGT 1620
Qy 544 eValArgAlaLeuGlyGlyAsnValSerLeuSerAspLeuAsnGlyGlyMetArgProGl 564
Db 1621 GTACGAGAGGCTAGAG-----AAGCTGCAGGCAGTGTGGCGGGGTG---CCCGG 1668
Qy 564 Y-----GlnSerAsnValTyrSerSer-- 571
Db 1669 GCATTTGGAGCGCCAGCTGCATCCCGCTTCCCGCAGGAGAACTCTACGTGTCCAG 1728
Qy 572 -----TyrGlySerThrAspTyrAsp----- 579
Db 1729 CACTGGCAGACCCAGCTGGGCTGCTCCATGACGCCCTGGCAGCCCATCAGGAGC 1788
Qy 580 -SerSerGlnTyrAsnGluAspMetLysLysPheArgLysMetAlaLeuGlyThrGlnGl 599
Db 1789 CAGTGGCCAGCAGCAGCAGCTGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1848
Qy 599 uTyrAsnAlaThrGlyGluTyrSerAsnProThrSerAspTyrGlyLeuTyrProSer-- 618
Db 1849 G-----AGCTAGCGCGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1902
Qy 619 -----GlySerSerSerGluGlyGlnThrThrArgGl 629
Db 1903 CCCTTGGACCCAGCACCCCTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1961
Qy 629 uMetGluMetGly-----LysileLysArgThrGlyGlnGlyTyrSerGlyPro 645
Db 1962 AATCAGCTGGGAGTGGCGCCAGGATCCCGGCCACAGCGCTGGAAGGACTGGCCCT 2019

RESULT 12

US-09-016-434-1093
; Sequence 1093, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Sellhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,434
; FILING DATE: HEREWITH
; CLASSIFICATION:


```

Db 1441 CACACTGCAAGCAGCTCTGGTGCAGATGCTTGGGCTGCTCCCATCGCCATGCAGATCTA 1500
QY 525 -----
Db 1501 CAAGAAGCACCTGGAGCCCGCCGCGCTGCTCCCATCGCTGAGCTGGGCGCTGGG 1560
QY 526 ---CysAlaAlaAlaCysValArgHisSerAlaArgArgProArgMetSerGlnI 544
Db 1561 CCAGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1620
QY 544 eValArgAlaLeuGluGlyAsnValSerLeuSerAspLeuAsnGluGlyMetArgProG 564
Db 1621 GTACGAGAGGCTAGAG-----AGCTGACGAGCAGTGTGTGGGGGTG---CCCGG 1668
QY 564 Y-----GlnSerAsnValTyrSerSer-- 571
Db 1669 GCATTGGAGCGCGCAGCTGCATCCCTCCCGCAGGAGAACTCTACGTGTCCAG 1728
QY 572 -----TyrGlyGlySerThrAspTyrAsp----- 579
Db 1729 CACTGGCAGAGCCACAGTGGGGTGTCTCCATGGCAGCCCTGGCAGCGCCATCAGGAGC 1788
QY 580 -SerSerGlnTyrAsnGluAspMetLysLysPheArgLysMetAlaLeuGlyThrGlnG 599
Db 1789 CAGTGGCCAGCAGCAGCAGCAGCTGCAGAGAGGCCCAACACGAGCCGCTGGAGAGTGACGA 1848
QY 599 uTyrAsnAlaThrGlyGluTyrSerAsnProThrSerAspTyrGlyLeuTyrProSer-- 618
Db 1849 G-----AGCTAGGCGGCTCTCTGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1902
QY 619 -----GlySerSerSerGluGlyGlnThrThrArgG 629
Db 1903 CCCTGTGACCCAGCACCCTCAGGAGGCGGCTGCTCCTCAGGGGACACGCG-AGGAG 1961
QY 629 uMetGluMetGly-----LysIleLysArgThrGlyGlnGlyTyrSerGlyPro 645
Db 1962 AATCAGCTGGGGAGTGGCCAGGATCCCGGCCACAGCGCTGGAGGAGTGGCCCT 2019

RESULT 13
PCT-US96-09193-1
; Sequence 1, Application PC/TUS9609193
; GENERAL INFORMATION:
; APPLICANT: Tularik, Inc.
; TITLE OF INVENTION: INTERLEUKIN-1 RECEPTOR-ASSOCIATED
; TITLE OF INVENTION: PROTEIN KINASE AND BINDING ASSAY
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/09193
; FILING DATE: JUNE 5 1996
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: U.S. Serial No. 08/587,889
; FILING DATE: JAN 16 1996
; CLASSIFICATION:
; APPLICATION NUMBER: U.S. Serial No. 08/494,006
; FILING DATE: JUNE 23 1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: David J. Brezner
; REGISTRATION NUMBER: 24,774
; REFERENCE/DOCKET NUMBER: FP-62191-1

```

```

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 494-8700
; TELEFAX: (415) 494-8771
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3590 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
PCT-US96-09193-1

Alignment Scores:
Pred. No.: 3,48e-16 Length: 3590
Score: 581.50 Matches: 208
Percent Similarity: 38.68% Conservative: 86
Best Local Similarity: 27.37% Mismatches: 239
Query Match: 16.84% Indels: 229
DB: 5 Gaps: 28

US-10-086-464-2 (1-647) x PCT-US96-09193-1 (1-3590)
QY 1 MetSerSerAlaProSerPro-GlyThrGlySerProSerProSerProSerAsnSerTh 20
Db 80 ATGGCCGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 139
QY 20 rThrThrThrProProAlaSerAla-----ProProPr 32
Db 140 GAGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 199
QY 32 oThrThrProSerProPro----- 39
Db 200 GACTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 259
QY 40 ----ProProSerThrThrProSerProProProSerProSerSerArgSerThrProSerAl 58
Db 260 CGCTCCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 319
QY 58 aProProProSerProPro-----ThrProSerTh 68
Db 320 GCCGACCTCGTGCACTCTCAGCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 379
QY 68 rProGlySerProProLeuProGlnProSerProProAlaProThrThrProGlySe 88
Db 380 GCCTGGCACCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 430
QY 88 rProProAlaProValThrProProThrArgAsnProProProSerValProGlyProPr 108
Db 431 CCCAGCAGCATCCCTGCGACCGCGCG-----AGGCGGAGGCGCTGGAGCCCGCG 479
QY 108 oSerAsnProSerArgGlyGlySerProArgProProProProSerSerProProSe 128
Db 480 -----AGTTGCCATCCTCGACGCTCCACCTTCCTTCCTC-TCGCCAGCTTT 519
QY 128 rProSerSerAspGlyLeuSerThrGlyValValValGlyLeuAlaIleGlyGlyValAl 148
Db 520 TCCAGGCTCCCGACCGCATTTCA----- 541
QY 148 aLeuLeuValIleValThrLeuIleCysLeuLeuCysLysLysLysArgArgAspG 168
Db 541 ----- 541
QY 168 uGluAspAlaTyrTyrValProProProProProProProGlyProLysAlaGlyGlyProTy 188
Db 542 -----GGGCGCTGAGCTCGGC----- 556
QY 188 rGlyGlyGlnGlnGlnTrpArgGlnGlnAsnAlaThrProProSerAspHisValVa 208
Db 556 ----- 556
QY 208 lThrSerLeuProProProLysAlaProSerProProProProProProProProProPr 228

```

Db 557 -----CTGGTTCCAGCCCTGCTTCCCTGTGGCTCCACCGCC 594
Qy 228 oProProPhoMetSerSerSerGlyGlySerAspTyrSerAspArgProValLeuPr 248
Db 595 ATCTCCAGCCCTTCTTCTACCAAGCCAGGC-----CC 627
Qy 248 oProProSerProGlyLeuValLeuGlyPheSerIysSerThrPheThrTyrGlu----- 266
Db 628 AGAGAGCTCAGTCCCTCCCTGCGAGGAGCCGCCCTCTCCGTTTGTGCTGGCCCTCTG 687
Qy 267 -GluLeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyGlyPh 286
Db 688 TGAGATTCCCGGGCACCACCACTTCTCGAGAGAGCTCAAGATCGGGAGGTGGCTT 747
Qy 286 eGlyTyrValHisIstysGlyValLeuProSerGlyLysGluValAlaValLysGlnLeuLy 306
Db 748 TGGGTGCTGTACCGGGCGGTGATG---AGGAACACCGGTGTATGCTGTGAAGAGCTGAA 804
Qy 306 sValGlySerGly-----GlnGlyGluArgGluPheGlnAlaGluValGluI 322
Db 805 GGAGAACCTGACCTGGAGTGCAGTGAAGCAGAGCTTCTCGACCGAGGTGGAGCA 864
Qy 322 eIleSerArgValHisArgHisLeuValSerLeuValGlyTyrCysIleAlaGlyAl 342
Db 865 CTTGTCAGGTTTCGTACCCACCAACTTGTGACTTTCCTGCTACTGTGCTCAGAACGG 924
Qy 342 aLysArgLeuLeuValTyrGluPheValProAsnAsnLeuGlu-----LeuHisLe 360
Db 925 CTTCTACTGCTGGTGTACGGCTTCTGCCCAACGGCTCCCTGGAGGACCGTCTCCACTG 984
Qy 360 uHisGlyGluGlyArgProThrMetGluTyrSerThrArgLeuLysIleAlaLeuGlySe 380
Db 985 CCAGACCCAGCCCTGCCACCTCTCTCTGCTGCTCAGCGACTGACATCTCTCTGGGTAC 1044
Qy 380 rAlaLysGlyLeuSerTyrLeuHisGluAspCysAsnProLysIleIleHisArgAspI 400
Db 1045 AGCCCGGCAATTGATTTTACATCAGGAC---AGCCCCAGCTTCATCCATGAGACAT 1101
Qy 400 eLysAlaSerAsnIleLeuIleAspPheLysPheGluAlaLysValAlaAspPheGlyLe 420
Db 1102 CAAGAGTTCCAAAGTCCTTCTGGATGAGAGGTGACACCCCAAGCTGGGAGACTTGGGCT 1161
Qy 420 uAlaLysIle-----AlaSerAspThrAsnThrHisValSerThrArg-- 434
Db 1162 GCGCGGTTACGCCCTTTGCGGGTCCAGCCCGCCAGCAGCAGCATGGTGGCCCGGAC 1221
Qy 435 -----ValMetGlyThrPheGlyTyrLeuAlaProGluTyrAlaAlaSerGlyLysLe 452
Db 1222 ACAGACAGTGGGGCACCTGGCTTCTGCTCCGAGGAGTACATCAAGACGGGAGGCT 1281
Qy 452 uThrGlyLysSerAspValPheSerPheGlyValValLeuLeuGluLeuIleThrGlyAr 472
Db 1282 GCGTGTGACACGGACACTTTCAGCTTTGGGTGGTGTAGTGTAGAGACTTGGCTGGTCA 1341
Qy 472 gArgProVal-----AspAlaAsnValTyrValAspAspSerLeuValAspTr 489
Db 1342 GAGGCTGTGAAGACGACCGGTGCGGAGCCCAAGTATCTGAAAGAC---CTGGTGGAA-- 1396
Qy 489 pAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGluGlyLeuAlaAspAl 509
Db 1397 -----GAGGAGGCTGAGAGGCTGGAGTGGCTTTGAGAAGCACCACGAG 1440
Qy 509 aLysMetAsnAsnGlyTyrAspArgGluGluMetAlaArgMetValAla----- 525
Db 1441 CACACTGCAAGCAGGTCTGGCTGCAGATGCTGGGCTGTCTCCATCGCCATGCAGATCTA 1500
Qy 525 ----- 525
Db 1501 CAAGAAGCAGCTGGACCCCGGCCCTTCCACCTGAGCTGGGCCCTGGGCTGGG 1560
Qy 526 ----CysAlaAlaAlaCysValArgHisSerAlaArgArgProArgMetSerGlnI 544
Db 1561 CCAGCTGGCTGCTGCTGCTGACCGCGGGGCCAAAGAGGCGCTCTCTATGACCCAGGT 1620

Qy 544 eValArgAlaIleuGluGlyAsnValSerLeuSerAspLeuAsnGluGlyMetArgProGl 564
Db 1621 GTACGAGAGGCTAGAG-----AAGCTGCAGGCAGTGTGCGGGGGTG---CCCGG 1668
Qy 564 Y-----GlnSerAsnValTyrSerSer-- 571
Db 1669 GCATTGGAGGCCGCCAGCTGCATCCCCCTTCCCCGAGGAGAACTCCTTACGTGTCCAG 1728
Qy 572 -----TyrGlyGlySerThrAspTyrAsp----- 579
Db 1729 CACTGGCAGAGCCACACAGTGGGCTGTCTCATGTCAGCCCTTGGCAGCGCCATCAGGAGC 1788
Qy 580 -SerSerGlnTyrAsnGluAspMetLysLysPheArgLysMetAlaLeuGlyThrGlnGl 599
Db 1789 CAGTGGCCAGCAGCAGCAGCAGCTGCAGAGAGGCCCCAACCCAGCCCCGTGGAGAGTGACGA 1848
Qy 599 uTyrAsnAlaThrGlyGluTyrSerAsnProThrSerAspTyrGlyLeuTyrProSer-- 618
Db 1849 G-----AGCTAGCGCGCTCTCTGCTGCGCTGGCTCTGGCCTTGACTTCAAGCTG 1902
Qy 619 -----GlySerSerSerGluGlyGlnThrThrArgGl 629
Db 1903 CCCTCTGGACCCAGCACCCCTCAGGAGGCGCGGTGTCTCAGGGGACACGGC-AGGAG 1961
Qy 629 uMetGluMetGly-----LysIleLysArgThrGlyGlnGlyTyrSerGlyPro 645
Db 1962 AATCAGCTGGGGAGTGGGCCAGGATCCCGGCCCCACAGCGGTGGAAGGACTGGCCCT 2019
RESULT 14
US-09-602-472A-5
; Sequence 5, Application US/09602472A
; Patent No. 6608240
; GENERAL INFORMATION:
; APPLICANT: Bidney, Dennis L.
; APPLICANT: Hu, Xu
; APPLICANT: Li, Guihua
; TITLE OF INVENTION: Sunflower Disease Resistance Genes
; FILE REFERENCE: 35718/200630
; CURRENT APPLICATION NUMBER: US/09/602,472A
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/140,876
; PRIOR FILING DATE: 1999-06-24
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 2114
; TYPE: DNA
; ORGANISM: Helianthus annuus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(2114)
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1475
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1475
; OTHER INFORMATION: n = A,T,C or G
; OTHER INFORMATION: n = A,T,C or G
US-09-602-472A-5

Alignment Scores:
Pred. No.: 2,96e-16 Length: 2114
Score: 578.00 Matches: 143
Percent Similarity: 55.07% Conservative: 85
Best Local Similarity: 34.54% Mismatches: 136
Query Match: 16.74% Indels: 51
DB: 4 Gaps: 15

US-10-086-464-2 (1-647) x US-09-602-472A-5 (1-2114)

US-08-265-628-1

Alignment Scores:

Pred. No.: 5,04e-16 Length: 2749
Score: 575.00 Matches: 154
Percent Similarity: 49.80% Conservative: 94
Best Local Similarity: 30.92% Mismatches: 160
Query Match: 16.65% Indels: 91
DB: 1 Gaps: 14

US-10-086-464-2 (1-647) x US-08-265-628-1 (1-2749)

```
Qy 130 SerSerAspGlyLeuSerThrGlyValValGlyLeuAlaIleGlyValAlaLeu 149
    ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Db 1327 AACCGAATGGAAACCATGACGCTGATTGTCGGAGTTTGTGTT----- 1371
Qy 150 LeuValIleValThrLeuIleCysLeuLeuCysValysLysArgArgAspGluGlu 169
    ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Db 1372 CTGCTCTTATGATCATGTTCTGC---CTCTGGAAAGGAAACAAAGCGAGCAAAACA 1428
Qy 170 AspAlaTyrValProProProProProGlyProLysAlaGlyGlyProTyrGly 189
    ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Db 1429 ACTGCAACATCTATTGTAAT----- 1449
Qy 190 GlyGlnGlnGlnGlnTrpArgGlnGlnAsnAlaThrProProSerAspHisValThr 209
    ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Db 1450 -----CGACAGAGAAAC-----CAAGATTTCGTAATGAAC 1479
Qy 210 SerLeuProProProProProProProProProProProProProProProPro 229
    ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Db 1480 GGGATG-----ATACTATCAAGCAAGACAGACAGTTCCT----- 1512
Qy 230 ProProPheMetSerSerSerGlySerAspTyrSerAspArgProValLeuProPro 249
    ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Db 1513 -----ATAGAGAACAAACCTGAGGAATTGGAACCTTCCATTGATA----- 1551
Qy 250 ProSerProGlyLeuValLeuGlyPheSerLysSerThrPheThrTyrGluGluLeuAla 269
    ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Db 1552 -----GAGTTTGGAAAGCTGTTGTC 1569
Qy 270 ArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyGlyPheTyrVal 289
    ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Db 1570 AAAGCCACCGAAATTTCTCAATTTGACAACTCGGCAAGGTGTTTCGGTATTGTT 1629
Qy 290 HisLysGlyValLeuProSerGlyLysGluValAlaValLysGlnLeuLysValGlySer 309
    ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Db 1630 TACAAGGTGATGATTACTTGTATGGGCAAGAAATTCGGTAAAAAGCTATCAAAAACGTCG 1689
Qy 310 GlyGlnGlyGluArgGluPheGlnAlaGluValGluIleLeuSerArgValHisArg 329
    ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Db 1690 GTTCAAGGGACTGGTGAGTTTATGAATGAGTGAGATTGATCGCGAGGCTTTCAGCATATA 1749
Qy 330 HisLeuValSerLeuValGlyTyrCysIleAlaGlyAlaLysArgLeuLeuValTyrGlu 349
    ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Db 1750 AACCTTGTCGAATTTCTGGCTGTTGATTGAGCAGCAGAGAGATGCTGTTATATGAG 1809
Qy 350 PheValProAsnAsnAsnLeuGluLeuHisLeuHisGlyGlyGluGlyArg---ProThrMet 368
    ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Db 1810 TATTAGAAAATTTAAGCTCGATCTTATCTCTTCGGAATAAAACGAAGCTCTACGTTA 1869
Qy 369 GluTrpSerThrArgLeuLysIleAlaLeuGlySerAlaLysGlyLeuSerTyrLeuHis 388
    ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Db 1870 AATTGGGAAGGACAGATTCAACATTACCAATGGTGTGCTCGAGACTTTTATATCTTCAT 1929
Qy 389 GluAspCysAsnProLysIleIleHisArgAspIleLysAlaSerAsnIleLeuIleAsp 408
    ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Db 1930 CAAGACTCACGGTTAGGATATCCACAGATATGAAGTAAGTAACATTTTCCTGAT 1989
Qy 409 PheLysPheGluAlaLysValAlaAspPheGlyLeuAlaLysIleAlaSerAspThrAsn 428
    ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Db 1990 AAAAATATGACACCAAGATCTCGGATTTTGGGATGCCAGAACTCTTTGCAAGGGACGAG 2049
Qy 429 ThrHisValSerThrArg---ValMetGlyThrPheGlyTyrLeuAlaProGluTyrAla 447
```

```
Db 2050 ACTGAAGCTAACACAGGAAGGTGGTCGGAACCTTACGGCTACATGTCTCCGGAGTAGCGCA 2109
Qy 448 AlaSerGlyLysLeuThrGluLysSerAspValPheSerPheGlyValValLeuLeuGlu 467
    ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Db 2110 ATGGATGGGGTATTCTCGGAAAATATCAGATGTTTTCAGTTCAGTTCATGTTCTTGAA 2169
Qy 468 LeuIleThrGlyArgArgProValAspAlaAsnValTyrValAspAspSerLeuVal 487
    ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Db 2170 ATTGTTAGTGAAAAAGGAACAGAGGATTCTACAACCTTGAACACCAAGAAAACATCTCTTA 2229
Qy 488 Asp-----TrpAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAsp----- 502
Db 2230 AGCTATATATGGAGTCTACGACGAGGAGGAGCGCTAGAAAATTGTTGATCCAGTCAATC 2289
Qy 503 PheGluGlyLeuAlaAspAlaLysMetAsnAsnGlyTyrAspArgGluGluMetAlaArg 522
    ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Db 2290 GTAGATTTCATTTGTCATCATTTACCAGCAACC-----TTTCAACCAAAAGAGTTCTAAAA 2343
Qy 523 MetValAlaCysAlaAlaAlaCysValArgHisSerAlaArgArgProArgMetSer 542
    ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Db 2344 TGCATACAAATTTGCTCTTGTGTGTTCAAGAACGTCGAGAGCATAGACCAACCATGTCG 2403
Qy 543 GlnIleValArgAlaLeuGluGlyAsnValSerLeuSerAspLeuAsnGluGlyMetArg 562
Db 2404 TCCGTGTTTGGATGCTTGGAAAGTGAAGCA-----ACAGAGATTCTCGAGCCTACACCG 2457
Qy 563 ProGlyGlnSerAsnValTyrSerSerTyrGlyGlySer-----ThrAspTyrAspSerSerGln 582
    ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Db 2458 CCAGTTTATCCCTCGGAAGAAGTCTCTTATGAAAATAATCTTTCATCAAGTAGACATTGC 2517
Qy 583 TyrAsnGluAspMetLysLysPheArgLysMetAlaLeuGlyThrGlnGluTyr 600
    ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
Db 2578 TACGA-AATCCGTTGAGAAAAGTTTCAGATAAATTAACTATTGGGGTGACCGGATAT 2630
```

Search completed: May 13, 2004, 01:36:26

Job time : 187 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 13, 2004, 00:33:06 ; Search time 677 Seconds
(without alignments)
4329.253 Million cell updates/sec

Title: US-10-086-464-2

Perfect score: 3453

Sequence:

1 MSSAPSGTGSPPSPNSNT.....REMEGKIKRTGGYSGPSL 647

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2941586 seqs, 2264995651 residues

Total number of hits satisfying chosen parameters: 5983172

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_epool_P/US10086464/runat_11052004_121817_7949/app_query.fasta_1.839
-DB=Published Applications NA -OFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOPFCI=0 -LOOPEXT=0 -UNITS=bites -START=1 -END=1 -MATRIX=blotsum62
-TRANS=human40.cdi -LIST=45 -DALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10086464 @CGN 1 1 333 @runat_11052004_121817_7949
-NCPU=6 -ICPU=3 -NO_WMAP -LARGESQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARG_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
------------	-------	-------------	--------	----	-------------

RESULT 1
US-10-086-464-1
; Sequence 1, Application US/10086464
; Publication No. US20020199218A1
; GENERAL INFORMATION:
; APPLICANT: GORING, Daphne R. et al.
; TITLE OF INVENTION: PROLINE-RICH EXTENSIN-LIKE RECEPTOR KINASES
; FILE REFERENCE: P 25,762-A USA
; CURRENT APPLICATION NUMBER: US 10/069,304
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 10/069,304
; PRIOR FILING DATE: 2002-02-19
; PRIOR APPLICATION NUMBER: PCT/CA00/009566
; PRIOR FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: US 60/149,466
; PRIOR FILING DATE: 1999-08-19
; PRIOR APPLICATION NUMBER: US 60/159,122
; PRIOR FILING DATE: 1999-10-13
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1944
; TYPE: DNA

ALIGNMENTS

1	3453	100.0	1944	14	US-10-086-464-1	Sequence 1, Appli
2	3453	100.0	2189	14	US-10-086-464-3	Sequence 3, Appli
3	2385.5	66.2	2451	13	US-10-424-599-95479	Sequence 95479, A
4	1870	54.2	1724	13	US-10-425-114-2326	Sequence 2326, Ap
5	1833	53.1	2880	13	US-10-425-114-33030	Sequence 33030, A
6	1833	53.1	2881	13	US-10-425-114-33031	Sequence 33031, A
7	1748.5	50.6	1902	14	US-10-086-464-10	Sequence 10, Appl
8	1748.5	50.6	1939	14	US-10-086-464-9	Sequence 9, Appli
9	1737.5	50.3	1424	13	US-10-425-114-9117	Sequence 9117, Ap
10	1687	48.9	2025	14	US-10-086-464-13	Sequence 13, Appl
11	1687	48.9	2104	14	US-10-086-464-12	Sequence 12, Appl
12	1539	44.6	2655	13	US-10-424-599-57846	Sequence 57846, A
13	1539	44.6	2655	13	US-10-425-114-7339	Sequence 7339, Ap
14	1437.5	41.6	2190	13	US-10-425-114-14127	Sequence 14127, A
15	1437	41.6	1500	13	US-10-425-114-6300	Sequence 6300, Ap
16	1437	41.6	1552	13	US-10-424-599-47702	Sequence 47702, A
17	1415.5	41.0	2196	14	US-10-086-464-7	Sequence 7, Appli
18	1415.5	41.0	2231	14	US-10-086-464-6	Sequence 6, Appli
19	1400.5	40.6	2261	14	US-10-086-464-15	Sequence 15, Appl
20	1399	40.5	2196	14	US-10-086-464-16	Sequence 16, Appl
21	1379	39.9	2508	13	US-10-425-114-25269	Sequence 25269, A
22	1335.5	38.7	1585	13	US-10-424-599-71008	Sequence 71008, A
23	1239	35.9	1588	13	US-10-425-114-35509	Sequence 35509, A
24	1238.5	35.9	1462	13	US-10-425-114-29328	Sequence 29328, A
25	1233.5	35.7	1134	13	US-10-425-114-36420	Sequence 36420, A
26	1210	35.0	1795	13	US-10-424-599-42232	Sequence 42232, A
27	1116.5	32.3	1179	13	US-10-425-114-4027	Sequence 4027, Ap
28	1116.5	32.3	1186	13	US-10-425-114-4273	Sequence 4273, Ap
29	1115.5	32.3	981	13	US-10-425-114-7549	Sequence 7549, Ap
30	921	26.7	729	13	US-10-425-114-12191	Sequence 12191, A
31	892.5	25.8	760	13	US-10-425-114-12314	Sequence 12314, A
32	866.5	25.1	951	13	US-10-425-114-1922	Sequence 1922, Ap
33	860.5	24.9	1777	13	US-10-425-114-24475	Sequence 24475, A
34	859.5	24.9	1817	13	US-10-425-114-30665	Sequence 30665, A
35	840	24.3	2025	13	US-10-425-114-3666	Sequence 3666, Ap
36	838.5	24.3	3148	16	US-10-260-238-465	Sequence 465, App
37	837	24.2	2284	13	US-10-425-114-3455	Sequence 3455, Ap
38	825	23.9	2012	13	US-10-424-599-29210	Sequence 29210, A
39	819.5	23.7	1011	13	US-10-425-114-13131	Sequence 13131, A
40	816.5	23.6	2124	9	US-09-938-842A-1079	Sequence 1079, Ap
41	816.5	23.6	2124	11	US-09-938-842A-1079	Sequence 311, App
42	811	23.5	3453	15	US-10-259-165-311	Sequence 25617, A
43	806.5	23.4	2063	13	US-10-424-599-25617	Sequence 33063, A
44	804	23.3	2374	13	US-10-425-114-33063	
45	802.5	23.2	2011	16	US-10-168-844-5	Sequence 5, Appli

```
; ORGANISM: Brassica napus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1944)
US-10-086-464-1

Alignment Scores:
Pred. No.:      9,36e-216      Length:      1944
Score:          3453.00      Matches:      647
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      100.00%      Indels:      0
DB:              14          Gaps:      0

US-10-086-464-2 (1-647) x US-10-086-464-1 (1-1944)

Qy 1 MetSerSerAlaProSerProGlyThrGlySerProProSerProProSerAsnSerThr 20
Db 1 ATGTCCTCGGCGCGCTCTCCGGGACTGGTTGGCTCCATCTCCACCATCAAAATCCACA 60

Qy 21 ThrThrProProProAlaSerAlaProProProProThrProSerSerProProPro 40
Db 61 ACCACCCTCTCTCCAGCTTCGCTCTCTCTCTCCACCACTCTCTCTCTCTCTCTCG 120

Qy 41 ProSerThrIleProThrSerProProProProProSerSerArgSerThrProSerAlaProPro 60
Db 121 CCATCCCACTATTCGGACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180

Qy 61 ProSerProProThrProSerThrProGlySerProProProProProGlnProSerPro 80
Db 181 CCATCTCCACCACTCCATCTACGCGGGAATCTCCACCTCTCTCTCTCTCTCTCTCTCA 240

Qy 81 ProAlaProThrThrProGlySerProProAlaProValThrProProProThrArgAsnPro 100
Db 241 CCGCTCCCACTACGCGCGGATCTCCACCGCACCTGTTACTCTCTCTCTCTCTCTCTCT 300

Qy 101 ProProSerValProGlyProProSerProSerAsnProSerArgGluGlyGlySerProArgPro 120
Db 301 CCACCTTCAGTCCGAGGACCACTCCCAATCTCTCCGCAAGGAGGATCTCTCTCGACCT 360

Qy 121 ProSerSerProSerProProSerProSerSerAspGlyLeuSerThrGlyValValVal 140
Db 361 CCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 420

Qy 141 GlyIleAlaIleGlyValAlaLeuLeuValIleValThrLeuIleCysLeuLeuCys 160
Db 421 GGAATCGCCATCGGAGGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 480

Qy 161 LysLysLysArgArgArgAspGluAspAlaTyrTyrValProProProProProProPro 180
Db 481 AAGAAGAAACACGCGAGAGACGAAAGATCTTACTATATGTTCTCTCGCCACCTCTCTCT 540

Qy 181 GlyProLysAlaGlyProTyrGlyGlyGlnGlnGlnGlnTyrAtgGlnGlnAsnAla 200
Db 541 GGTCCCAAGCCGAGGACCTTACGGTGGACAGCAGCAACAATGGCGCAACAACACGCA 600

Qy 201 ThrProProSerAspHisValValThrSerLeuProProProProProLysAlaProSerPro 220
Db 601 ACACCACGTCAGATCATGTCTGTCGTCAGTCACTTACCACCACTTAAGGCTCCATCTCCA 660

Qy 221 ProArgGlnProProProProProProProProProPheMetSerSerSerGlyGlySerAsp 240
Db 661 CCACGGCAACCTCTCTCACTCCACCTCCACCGCTTTCATGACGACGCGCGGCTCCGAC 720

Qy 241 TyrSerAspArgProValLeuProProProProProProGlyLeuValLeuGlyPheSerLys 260
Db 721 TACTCGGACCGCTCAGTTCTCTCTCCACCGTCTCCAGGCTTGTGTAGGCTTCTCCAA 780

Qy 261 SerThrPheThrTyrGluGluLeuAlaAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeu 280
Db 781 AGCACTTTCAATACGAGGAGTAGTAGAGCCACCAATGTTTCTCCGAGGCGCAACTTG 840

Qy 281 LeuGlyGlnGlyPheGlyTyrValHisLysGlyValLeuProSerGlyLysGluVal 300

841 TTAGGACAAAGCGGGTTCGGTTACGTGCACAAAGGTGTGTTCCTAGTGGGAAGAAGTT 900
Qy 301 AlaValLysGlnLeuLysValGlySerGlyGlnGlyGluArgGluPheGlnAlaGluVal 320
Db 901 GCTGTGAAGCAGTTGAAGTTGGAGTGGTCAGGAGAGAGGAGTTTCAGCGAGAGGTT 960

Qy 321 GluIleSerArgValHisHisArgHisLeuValSerLeuValGlyTyrCysIleAla 340
Db 961 GAGATCATCAGCAGATTCCACACAGGCATCTGGTGTCTCTTGTGTGGTTATTGTCATCGCC 1020

Qy 341 GlyAlaLysArgLeuLeuValTyrGluPheValProAsnAsnLeuLeuGluHisLeu 360
Db 1021 GGTGCCAAAAGATTGCTTGTCTATGAGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1080

Qy 361 HisGlyGluGlyArgProThrMetGluTrpSerThrArgLeuLysIleAlaLeuGlySer 380
Db 1081 CATGGCGAGGACGGCCCTACATGGAATGGAGCACCAGATTGAAGATTGCTCTTGGATCT 1140

Qy 381 AlalysGlyLeuSerTyrLeuHisGluAspCysAsnProLysIleIleHisArgAspIle 400
Db 1141 GCTAAAGGACTTCTTATCTTCAAGAGATTGCAATCTTAAATCAATCAATCAATCAATCA 1200

Qy 401 LysAlaSerAsnIleLeuIleAspPheLysPheGluAlaLysValAlaAspPheGlyLeu 420
Db 1201 AAGGCTTCAACACATATTGATAGATTTCAGTTTGAAGCTAAGGTTGCTGATTTTGGTCTT 1260

Qy 421 AlalysIleAlaSerAspThrAsnThrHisValSerThrArgValMetGlyThrPheGly 440
Db 1261 GCTAAGATTGCTTCTGATACAAACACGCATCTATCAACCGTGTGATGGGAACCTTTGGG 1320

Qy 441 TyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPheSer 460
Db 1321 TACTTGGCTCCGGAATACGCTCAAGCGGAAGCTCAGCGAGAAGTCTGACGTTTCTCA 1380

Qy 461 PheGlyValValLeuLeuLeuLeuLeuThrGlyArgArgProValAspAlaAsnVal 480
Db 1381 TTTGGCGTGTGCTTTTGGAGCTCATTTACTGGAGCTGACCCGTTGATGCCAACAATGTC 1440

Qy 481 TyrValAspSerLeuValAspTrpAlaArgProLeuLeuAsnArgAlaSerGluGln 500
Db 1441 TATGTAGATGACAGCTTAGTTGACTGGGCACGACCATTTGCTTAACCGAGCATCTGAGCA 1500

Qy 501 GlyAspPheGluGlyLeuAlaAspAlaLysMetAsnAsnGlyTyrAspArgGluGluMet 520
Db 1501 GGAGACTTTGAGGGTTTAGCTGTATGCAAGATGAATAATGGGTATGACAGAGAGAGATG 1560

Qy 521 AlaArgMetValAlaCysAlaAlaAlaCysValArgHisSerAlaArgArgProArg 540
Db 1561 GCTCGCATGGTGTCTGTGCTGCGGCTTGTGTTCCCATTTGAGCTCGCCGAGACCTTCG 1620

Qy 541 MetSerGlnIleValArgAlaLeuGluGlyAsnValSerLeuSerAspLeuAsnGluGly 560
Db 1621 ATGAGCCAGATTGCGCGTGGTTAGAAAGGAAATGTATCACTGTCTGATCTTAAACGAAGG 1680

Qy 561 MetArgProGlyGlnSerAsnValTyrSerSerTyrGlyGlySerThrAspTyrAspSer 580
Db 1681 ATGAGACAGGTCAAAGCAATGTATACAGCTCATACGAGGAGGAGCAACCGATTATGACTCG 1740

Qy 581 SerGlnTyrAsnGluAspMetLysPheArgLysMetAlaLeuGlyThrGlnGluTyr 600
Db 1741 AGCAGTACATGAATGAACATGAAGAGTTTAGGAAATTTGGCACTTGGCACTCAAGAGTAC 1800

Qy 601 AsnAlaThrGlyGlyTyrSerAsnProThrSerAspTyrGlyLeuTyrProSerGlySer 620
Db 1801 AACGCCACGGGTAGTACAGTAAATCCGACCATGACTATGAGCTGTATGAGCTCTCTGTTCA 1860

Qy 621 SerSerGluGlyGlnThrThrArgGluMetGluMetGlyLysIleLysArgThrGlyGln 640
Db 1861 AGCAGCGAGGCGCAACCAACACCGGAAATGGAGATGGGAAGATTAAGAGAACCCTGTCAG 1920

Qy 641 GlyTyrSerGlyProSerLeu 647
```

Db 1921 GGTATAGTGGACCTTCTT 1941

RESULT 2

US-10-086-464-3

; Sequence 3, Application US/10086464

; Publication No. US20020199218A1

; GENERAL INFORMATION:

; APPLICANT: GORING, Daphne R. et al.

; TITLE OF INVENTION: PROLINE-RICH EXTENSIN-LIKE RECEPTOR KINASES

; FILE REFERENCE: P 25,762-A USA

; CURRENT APPLICATION NUMBER: US/10/086,464

; CURRENT FILING DATE: 2002-02-28

; PRIOR APPLICATION NUMBER: US 10/069,304

; PRIOR FILING DATE: 2002-02-19

; PRIOR APPLICATION NUMBER: PCT/CA00/00966

; PRIOR FILING DATE: 2000-08-18

; PRIOR APPLICATION NUMBER: US 60/149,466

; PRIOR FILING DATE: 1999-08-19

; PRIOR APPLICATION NUMBER: US 60/159,122

; PRIOR FILING DATE: 1999-10-13

; NUMBER OF SEQ ID NOS: 27

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 3

; LENGTH: 2189

; TYPE: DNA

; ORGANISM: Brassica napus

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (1)..(2189)

; OTHER INFORMATION:

US-10-086-464-3

Alignment Scores:

Pred. No.: 1,05e-215 Length: 2189

Score: 3453.00 Matches: 647

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 14 Gaps: 0

US-10-086-464-2 (1-647) x US-10-086-464-3 (1-2189)

Qy 1 MetSerSerAlaProSerProGlyThrGlySerProProSerProProSerAsnSerThr 20

Db 97 ATGTCCTCGCGCGCTCTCGGGGACTGGTTCGCTCCATCTCCACCATCAAACTCCACA 156

Qy 21 ThrThrThrProProAlaSerAlaProProProThrProSerProProPro 40

Db 157 ACCACCACTCTCTCCAGCTTCGGCTCTCTCTCCACCAACACACCTCTCTCTCTCCGCG 216

Qy 41 ProSerThrThrProThrSerProProProSerSerSerArgSerThrProSerAlaProPro 60

Db 217 CCATCCCACTATTCGACATCT 276

Qy 61 ProSerProProThrProSerThrProGlySerProProProProGlnProSerPro 80

Db 277 CCATCTCCCACTCCATCTACCGCGGATCTCCACTCTCTCTCTCTCTCTCTCTCTCTCTCT 336

Qy 81 ProAlaProThrThrProGlySerProProAlaProValThrProProThrArgAsnPro 100

Db 337 CCGCTCCAACTACGCGCGGATCTCCACCGCGACCTGTACTCTCTCTCTCTCTCTCTCTCTCT 396

Qy 101 ProProSerValProGlyProProProSerAsnProSerArgGluGlyGlySerProArgPro 120

Db 397 CCACCTTCAGTCCCGGACCAACCGTCCCAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 456

Qy 121 ProSerSerProSerProProSerProSerSerAspGlyLeuSerThrGlyValValVal 140

Db 457 CCATCT 516

Qy 141 GlyIleAlaIleGlyValAlaLeuLeuValIleValThrLeuIleCysLeuLeuCys 160

Db 517 GGAATCGCCATCGGAGGAGTCT 576

Qy 161 LysLysLysArgArgArgAspGluGluAspAlaTyrTyrValProProProProProPro 180

Db 577 AAGAGNAACGACGGAGAGACGAGAGATGCTTACTATGTTCTCTCGCCACCTCTCTCTCT 636

Qy 181 GlyProLysAlaGlyGlyProTyrGlyGlyGlnGlnGlnGlnTrpArgGlnGlnAsnAla 200

Db 637 GGTCCCAAGCCGAGGACCTTACGTTGGGAGCAGCAGCAACAATGGCGGCAACAAGCGCA 696

Qy 201 ThrProProSerAspHisValValThrSerLeuProProProProProProProProPro 220

Db 697 ACACCAACCGTCAGATCATGTCGTGACGTCACTACCACCACTTAAGGCTTCCATCTCCA 756

Qy 221 ProArgGlnProProProProProProPheMetSerSerSerSerGlyGlySerAsp 240

Db 757 CCACGGCAACTCTCTCCACCTCCACCCGCTTTCATGAGCAGCAGCGGCGGCTCCGAC 816

Qy 241 TyrSerAspArgProValLeuProProProSerProGlyLeuValLeuGlyPheSerLys 260

Db 817 TACTCGGACCGTCCAGTTCTCTCCACCGCTCTCCAGGCTTGTGTAGGCTTCTCCAAA 876

Qy 261 SerThrPheThrTyrGluGluLeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeu 280

Db 877 AGCACTTTCACATACGAGGAGCTAGCTAGAGCCACCAATGCTTCTCCGAGGCGAAGTTG 936

Qy 281 LeuGlyGlnGlyGlyPheGlyTyrValHisLysGlyValLeuProSerGlyLysGluVal 300

Db 937 TTAGGACNAGCGGGTTCGGTTAGTGACCAAAAGTGTGTGCTAGTGGGAAGAAGTT 996

Qy 301 AlaValLysGlnLeuLysValGlySerGlyGlnGlyGluArgGluPheGlnAlaGluVal 320

Db 997 GCTGTGAAGCAGTTGAAAGTTGGAGTGGTTCAGGAGAGGAGGAGTTTTCAGGAGAGGTT 1056

Qy 321 GluIleSerArgValHisArgHisLysValSerLeuValGlyTyrCysIleAla 340

Db 1057 GAGATCATATCAGAGAGTTCCACAGGACATCTGGTGTCTCTGTGTGTATTTGTCATCGCC 1116

Qy 341 GlyAlaLysArgLeuLeuValTyrGluPheValProAsnAsnAsnLeuGluLeuHisLeu 360

Db 1117 GGTGCCAAAAGATTTGTTGTTCTATGAGTTTGTCTTAAACAACAATCTCGAGCTTCACTC 1176

Qy 361 HisGlyGluGlyArgProThrMetGluTrpSerThrArgLeuLysIleAlaLeuGlySer 380

Db 1177 CATGGCGAGGACGGCCCTACAATGGAATGGAGACACCAAGATTGAAGATTGCTCTTGGATCT 1236

Qy 381 AlaLysGlyLeuSerTyrLeuHisGluAspCysAsnProLysIleIleHisArgAspIle 400

Db 1237 GCTAAAGGAGCTTTCTTATCTTCAAGATTGCAATCTTAAATCATTCACCGGTGATATC 1296

Qy 401 LysAlaSerAsnIleLeuIleAspPheLysPheGluAlaLysValAlaAspPheGlyLeu 420

Db 1297 AAGCTTCAAAACATATTGATAGATTTCAGTTTGAAGCTTGAAGCTTGTGATTTTGTGCTT 1356

Qy 421 AlaLysIleAlaSerAspThrAsnThrHisValSerThrArgValMetGlyThrPheGly 440

Db 1357 GCTAAGATTCTCTCTGATACAAACACGATGATCAACAGTGTGATGGGAACCTTTTGGG 1416

Qy 441 TyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPheSer 460

Db 1417 TACTTGGCTCCGGAATACGCTGCAAGCGGAAGCTCACGGAAGTCTGACGTTTCTCA 1476

Qy 461 PheGlyValValLeuLeuGluLeuIleThrGlyArgArgProValAspAlaAsnAsnVal 480

Db 1477 TTTGGGCTGTGCTTTTGGAGCTCATTTACTTGGAGCTCGACCGCTTGTATGCCAACATATGC 1536

Qy 481 TyrValAspAspSerLeuValAspTrpAlaArgProLeuLeuAsnArgAlaSerGluGln 500

Db 1537 TATGTAGATCACAGCTTAGTTGACTGGGCACGACCATTTGCTTAACCGAGCATCTGACCAA 1596

Qy 501 GlyAspPheGluGlyLeuAlaAspAlaLysMetAsnAsnGlyTyrAspArgGluGluMet 520

Db 1597 GGAGACTTTGAGGCTTGTGATGCAAAAGATGAATAATGGTATGTAGCAGAGAGAGATG 1656

```
Qy 521 AlaArgMetValAlaCysAlaAlaCysValArgHisSerAlaArgArgProArg 540
Db 1657 GCTCGCATGGTGTGTGCTGCGGCTGTGTTCGCCAATTCAGCTCGCGCAGACCTCGC 1716
Qy 541 MetSerGlnIleValArgAlaLeuGluGlyAsnValSerLeuSerAspLeuAsnGluGly 560
Db 1717 ATGAGCCAGATTGTCGTCGCTTAGAGGAATGTATCATCTGTAGATCTTAAAGAGGG 1776
Qy 561 MetArgProGlyGlnSerAsnValTyrSerSerTyrGlyGlySerThrAspTyrAspSer 580
Db 1777 ATGAGCAGGTCAAAGCAATGTATACAGCTCATACGAGGAAGCACCAGATTATGACTCG 1836
Qy 581 SerGlnTyrAsnGluAspMetLysLysPheArgLysMetAlaLeuGlyThrGlnGluTyr 600
Db 1837 AGCCAGTACAAATGAAGACATGAAGAAGTTTAGGAAAATGGCACATTGGAACCTCAAGAGTAC 1896
Qy 601 AsnAlaThrGlyGluTyrSerAsnProThrSerAspTyrGlyLeuTyrProSerGlySer 620
Db 1897 AACGCCACGGGTGAGTACAGTAATCCGACCACTGACTATGGACTGTACCCGCTGTTTCA 1956
Qy 621 SerSerGluGlyGlnThrThrArgGluMetGluMetGlyLysLysLysArgThrGlyGln 640
Db 1957 AGCAGCGAGGCGCAACACACACGCAATGGAGATGGGAAGATTAAAGAAACCGGTTCAG 2016
Qy 641 GlyTyrSerGlyProSerLeu 647
Db 2017 GGTATTAGTGACCTTCTCTT 2037

RESULT 3
US-10-424-599-95479
; Sequence 95479, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 95479
; LENGTH: 2451
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_57229C.1
US-10-424-599-95479

Alignment Scores:
Pred. No.: 9.23e-140 Length: 2451
Score: 2285.50 Matches: 450
Percent Similarity: 80.86% Conservative: 74
Best Local Similarity: 69.44% Mismatches: 97
Query Match: 66.19% Indels: 27
DB: 13 Gaps: 16

US-10-086-464-2 (1-647) x US-10-424-599-95479 (1-2451)

Qy 13 ProSerProProSerAsnSerThrThrThrThrProProProAla-----SerAlaPro 30
Db 246 CCCCTCCCTCGGCTC---GCAACTCCGTCGTCACCTCGCGCGGAACCTCCGCTGCTCCT 302
Qy 31 ProProThrThrProSerProProProSerThrThrThrProProProProPro 50
Db 303 CCGCGCGGACACCTTCTTCGCCACCTCCG-----TCAACTCTCTTCCTCGCCCTCCG 356
Qy 51 SerSerArgSerThrProSerAlaProProProSerProProProSerThrProGly 70
Db 357 -----GGACTCCCTCGGCTTCTCCACCGTCCACTCTCTCTGCTTCTCTCTCCACCG 407
```

```
Qy 71 SerProProProLeuProGlnProProSerProProAlaProThrThrProGlySerProPro 90
Db 408 TCCACTCAACACCGCGCTCAACTTTCGCCGCGCATCG---ACTTCTCCGCGCTGCCCGCA 464
Qy 91 -----AlaProValThrProProThrArgAsnProProProProSerValProGlyProPro 108
Db 465 TCGCACTCGCGCGCTCGCTCCAGTGGCGGCGGAGCGCAGCACTCCGAGTCCACCG 524
Qy 109 SerAsnProSerArgGluGlyGlySerProArgProProSerSerProSerProSerProSer 128
Db 525 TCCGGAGCTCGGCTCTCCTCGCTCGGATTCGAGACCGACCACTCCGCTCGCTCCGTCG 584
Qy 129 ProSerSerAspGlyLeuSerThrGlyValValGlyLysLysLysLysLysLysLysLys 148
Db 585 TCTTCTCTCGTCGAGTATTCGACCGGTGTGTGTGGGAATCGCGGTGGGGCTGTGGCG 644
Qy 149 LeuLeuValIleValThrLeuIleCysLeuLeuCys---LysLysLysLysLysLysLysLys 167
Db 645 GTTCTTCTGTGTGAGCATTCCTGCAATGTTGCGGGAAGAAGAAGAGACGCTGAT 704
Qy 168 GluGluAspAlaTyrTyrValProProProProProProProProProProProProProPro 186
Db 705 GAAGAG-----TACTATGCTCGCGCGCGCAACCGCGCGGCGGACCTAAA---GATGAT 755
Qy 187 ProTyrGlyGlyGlnGlnGlnTyrArgGlnGlnAsnAlaThrProProSerAspHis 206
Db 756 GCATATGCTGCTCCCGCACGTCATATG---CAACACAAATGTTCCCTCTCTCAAGATCAT 812
Qy 207 ValValThrSerLeuProPro-----ProProLysAlaProSerProProProProPro 223
Db 813 GTGGTCTCAATGATGCTCCAAAGCCATCGCCACCACTGCTCCACCGGCTTATGCTGGT 872
Qy 224 ProProProProProProProProProPheMetSerSerSerGly-----GlySerAspTyr 241
Db 873 CAACCTCCCCACCAACCGCTCTCTTTCATCAGCAGCAGTGGCGGATCTGGATCAAACTAT 932
Qy 242 SerAspArgProValLeuProProProProGlyLeuValLeuGlyPheSerLysSer 261
Db 933 TCAGCGCGTGAAATTTCTCTCTCTCTCTCCAGAAATTTCAATGGGATTTCTTAAAGAGC 992
Qy 262 ThrPheThrTyrGluGluLeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeu 281
Db 993 ACATTACGATGAGGAGTTGGCAGCGCAACTGATGCTCTCTGATGCCCACTCCTT 1052
Qy 282 GlyGlnGlyGlyPheGlyTyrValHisLysGlyValLeuProSerGlyLysGluValAla 301
Db 1053 GGACAGAGGAGATTGGATATGTGCACAGAGGAATTTCTCCCAACGGCAGGAGGTGGCA 1112
Qy 302 ValLysGlnLeuLysValGlySerGlyGlnGluArgGluPheGlnAlaGluValGlu 321
Db 1113 GTGAAGCAATTGAAGGCTGGAAGCGGCGCAAGGGAGCGTGAATTCCAAGCTGAAGTTGAG 1172
Qy 322 IleIleSerArgValHisHisArgHisLeuValSerLeuValGlyTyrCysIleAlaGly 341
Db 1173 ATAATTAGCGCTGTCATCACAGCATCTGTTTCTTTGGTTGGTATCTGATCATCATCGG 1232
Qy 342 AlaLysArgLeuLeuValTyrGluPheValProAsnAsnAsnLeuGluLeuHisLeuHis 361
Db 1233 TCCAGAGGCTGCTTGTATGAATTTGTTTCCCAACACACATTTGGAATTTCCATTTGCGAT 1292
Qy 362 GlyGluGlyArgProThrMetGluTyrSerThrArgLeuLysLysLysLysLysLysLys 381
Db 1293 GGAAGAGGCGGACCTACCTGAGTTGGCCCAAGATTCATCTTAAGATCATCTCGTATCAAA 1352
Qy 382 LysGlyLeuSerTyrLeuHisGluAspCysAsnProLysLysLysLysLysLysLysLys 401
Db 1353 AAGGACTGGCGTATCTTCATGAGATTGTCATCTTAAGATCATCTCGTATCAAA 1412
Qy 402 AlaSerAsnLeuLeuIleAspPhePheGluAlaLysValAlaAspPheGlyLeuAla 421
Db 1413 GCTGCCAACATCTCTCGGATTTTAAGTTTGAAGCAAGGTTGCGAGATTTTGGTCTTGCA 1472
Qy 422 LysIleAlaSerAspThrAsnThrHisValSerThrArgValMetGlyThrPheGlyTyr 441
```

```
Db 1473 AAGTTTCTCTGATGCAATACCCATGTTTCTACTCGAGTGATGGGACTTTTGGGTAT 1532
Qy 442 LeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPheSerPhe 461
Db 1533 TTGGCTCCAGAAATATGCTTCTAGTGGAAACTGACAGCAAAATCAGATGTTTCTCTCTAT 1592
Qy 462 GlyValValLeuLeuGluLeuThrGluValArgProValAspAlaAsnAsnValTyr 481
Db 1593 GGAGTCATGCTCTCGAGTTAATACCGGACGCGCGGCTCGATTAATAATCAAACTTTC 1652
Qy 482 ValAspSerLeuValAspTTPAlaArgProLeuLeuAsnArgAlaSerGluGlnGly 501
Db 1653 ATGGAGGATAGTTGGTAGACTGGGCTAGGCTTGTCTCACAGGCTTTGGAAAGAGAT 1712
Qy 502 AspPheGluGlyLeuAlaAspAlaLysMetAsnAsnGlyTyrAspArgGluGluMetAla 521
Db 1713 GATTTTGATCTATATTGACCCCAAGGCTCCAGAAATGACTATGATCTTAATGAGATGGCA 1772
Qy 522 ArgMetValAlaCysAlaAlaCysValArgHisSerAlaArgArgArgProArgMet 541
Db 1773 CGAATGGTGGCTTCTGTGGGCTTGCAATTCGTCTATTCGGCAAGCGTCGACCAAGGATG 1832
Qy 542 SerGlnLeuValArgAlaLeuGluGlyAsnValSerLeuSerAspLeuAsnGluGlyMet 561
Db 1833 AGCCAGGTGTTGCGGCTCTGGAAGGAGATGCTCTCTAGCAGATCTTAACGAAGGAATT 1892
Qy 562 ArgProGlyGlnSerAsnValTyrSerSerTyrGlyGlySerThrAspTyrAspSerSer 581
Db 1893 AGACCTGGACACACACTATACAGTTCTCAT---GAAAGCTCAGATTAATGACACATGCA 1949
Qy 582 GlnTyrAsnGluAspMetLysPheArgLysMetAlaLeuGlyThrGlnGluTyrAsn 601
Db 1950 CAGTACGAAGGACATGAAAGTTTCAAGAAATGGCAATGGCAATCTCAGGAGTATGGT 2009
Qy 602 AlaThrGlyGluTyrSerAsnProThrSerAspTyrGlyLeuTyrProSerGlySerSer 621
Db 2010 GCAAGCAGTGAGTACAGTCCGCTACAAGTGAGTATGGTTTAAACCCATCAGGCTCAAGT 2069
Qy 622 SerGluGly-----GlnThrArgGluMetGluMetGlyLysIleLysArgThr 638
Db 2070 AGTGAACACACAGACCGCCCAACCAAGGGAATGGAAATGGAAGATGAAG---AAC 2126
Qy 639 GlyGlnGlyTyrSerGlyProSer 646
Db 2127 AATCAAGTTTTCAGTGAAGTTCT 2150
```

RESULT 4

```
US-10-425-114-2326
; Sequence 2326, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425.114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 2326
; LENGTH: 1724
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700209610_FU1
US-10-425-114-2326
Alignment Scores:
```

```
Pred. No.: 6,76e-113 Length: 1724
Score: 1870.00 Matches: 363
Percent Similarity: 79.85% Conservative: 53
Best Local Similarity: 69.67% Mismatches: 87
Query Match: 54.16% Indels: 18
DB: 13 Gaps: 9
US-10-086-464-2 (1-647) x US-10-425-114-2326 (1-1724)
Qy 139 ValValGlyLysLeuAlaLysGlyValAlaLeuLeuValLysLeuLeuValLysLeu 158
Db 1 GTGCTGGCGTGGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 60
Qy 159 LeuCys-----LysLysLysArgArgAspGluGluAspAlaLysValProPro 176
Db 61 TGCTGTCTCCGCAAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
Qy 177 ProProProGlyProLysAlaGlyGlyProTyrGlyGlyGlnGlnGlnGlnTyrArg 196
Db 121 CCGCGCGCGCGCG---CCGTACAAGAGGATCCATACGCTGGAACGTACCAAGATTGG--- 174
Qy 197 GlnGlnAsnAla-----ThrProProSerAspHisValValThrSerLeuPro---Pro 213
Db 175 CAGCAAAATCGGCTCTGCTCCACCCCTGCAACATGTGTCAAGATGCACCTTCGCGCT 234
Qy 214 ProProLysAlaProSerProArgGlnProProProProProProProProProPro 233
Db 235 CCGCGCAGCATATGCCAATCGCTCTCCACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCTA 294
Qy 234 SerSerSerGly-----GlySerAspTyrSerAspArgProValLeuLeuProProSer 251
Db 295 AATAGTAGTGGTGGATCTGCTTCTAATTAATCTGCTGGCGAGATCTTACCTCCACCATCC 354
Qy 252 ProGlyLeuValLeuGlyPheSerLysSerThrPheThrTyrGluGluLeuAlaArgAla 271
Db 355 CCGTGGTCTCTTCTGGCTTCTCGAAGAGCACAATTCACCTACGAAGAGTGTGTCAGGCGCT 414
Qy 272 ThrAsnGlyPheSerGluAlaAsnLeuGlyGlnGlyGlyPheGlyTyrValHisLys 291
Db 415 ACTGATGGATCTCGGATGCTATCTCTTGGACAGAGTGGTTTGGCTATGTTACAGA 474
Qy 292 GlyValLeuProSerGlyLysGluValAlaValLysGlnLeuLysValGlySerGlyGln 311
Db 475 GGATTTGCTGCTAATGGCAAGAGATTGCTGTAAAAACAACCTGAAACTGGGAAGTGGCCAG 534
Qy 312 GlyGluArgGluPheGlnAlaGluValGluLysSerArgValHisHisArgHisLeu 331
Db 535 GCGGAGCGTGAGTTCAGGCTGAGTTGAGATTATCAGCGGAGTACATCAAAACACCTT 594
Qy 332 ValSerLeuValGlyTyrCysIleAlaGlyAlaLysArgLeuLeuValTyrGluPheVal 351
Db 595 GTGCTTTGGTTGGCTATTGCAATTTCTGGAGGCAAGAGGCTGCTTGTCTATGAGTTTGT 654
Qy 352 ProAsnAsnLeuLeuGluLeuHisLeuHisGlyGluGlyArgProThrMetGluTyrSer 371
Db 655 CCCAATAACATTTGGAATTCCTACTTACATCGCAAGATCGAACCAACAATGGAGTGGCCT 714
Qy 372 ThrArgLeuLysIleAlaLeuGlySerAlaLysGlyLeuSerTyrLeuHisGluAspCys 391
Db 715 GCTAGATTAAAGATCAGTTTGGGTGCTGCCAAGGGTTAGCTTATCTTCTATGAAGACTGC 774
Qy 392 AsnProLysIleIleHisArgAspIleLysAlaSerAsnIleLeuIleAspPheLysPhe 411
Db 775 CATCCAAAGATCATCCATCTGATCAATGAAGCATCTAACATCTTCTTGTGACTTCAATTT 834
Qy 412 GluAlaLysValAlaAspPheGlyLeuAlaLysIleAlaSerAspThrAsnThrHisVal 431
Db 835 GAAGCTAAGGTGCTGATTTTGGACTTTGCAAGTTTCACTACTGATAACAACACCATGTT 894
Qy 432 SerThrArgValMetGlyThrPheGlyTyrLeuAlaProGluTyrAlaAlaSerGlyLys 451
Db 895 TCGACAAGAGTAATGGGCACCTTTGGGTATTTGGCACCTTGGTATGATGATCTTCTGGCAAG 954
```


2213	TATGATGCTTTAGTGGATCTCTCGCTGGGAAAGGACTTCAATCTTAATGAGATGGCAAGA	2272
Qy	523 MetValAlaCysAlaIalaAlaCysValArgHisSerAlaAlaArgArgProArgMetSer	542
Db	2273 ATGATAGCCTGTGCAGCTGTCATGTGTACGCCATTTCTGCACGTCGTGCGCCACGTTATGAGT	2332
Qy	543 GlnIleValArgAlaLeuGluGlyAsnValSerLeuSerAspLeuAsnGluGlyMetArg	562
Db	2333 CAGGTCTTTCGGGCTTTGGAGGGCAATGTGTCTTTGGAGGACCTTAATGAAGGTGTTTCGG	2392
Qy	563 ProGlyGlnSerAsnValTyrSerSerTyrGlyGlySerThrAspTyrAspSerSerGln	582
Db	2393 CCTGGCCATAGCCGCTTCTTTGGGTCAATAC---AGCAGCTCCGATTACGATTTCTGGCCAG	2449
Qy	583 TyrAsnGluAspMetLysLysPheArgLysMetAlaLeuGlyThrGlnGluTyrAsnAla	602
Db	2450 TACAACGAGGACATGAAGAAGTTTCAAGAAGATGGCAATTCACAACAAC-----TATACC	2503
Qy	603 ThrGlyGluTyrSerAsnProThrSerAspTyrGlyLeuTyrProSerGlySerSerSer	622
Db	2504 AGCAGCCAATACACGCGGCCAACAGTGAATATATGCACAGATACCGTCTGCATCAAGCAGC	2563
Qy	623 GluGlyGlnThrThrArgGluMetGluMetGlyLysIleLysArgThrGly-----	639
Db	2564 GAGGCCACCACGACGAGGAGATGGATCGGTGCATGAAGAAGGTGGCTACAGTGGT	2623
Qy	640 -----GlnGlyTyrSerGlyProSer	646
Db	2624 GGCTACAGCTCAGGATACAGCGGAGCCTCG	2653

RESULT 7

```

US-10-086-464-10
; Sequence 10, Application US/10086464
; Publication No. US20020199218A1
; GENERAL INFORMATION:
; APPLICANT: GORING, Daphne R. et al.
; TITLE OF INVENTION: PROLINE-RICH EXTENSIN-LIKE RECEPTOR KINASES
; FILE REFERENCE: P 25,762-A USA
; CURRENT APPLICATION NUMBER: US/10/086,464
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 10/069,304
; PRIOR FILING DATE: 2002-02-19
; PRIOR APPLICATION NUMBER: PCT/CA00/00966
; PRIOR FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: US 60/149,466
; PRIOR FILING DATE: 1999-08-19
; PRIOR APPLICATION NUMBER: US 60/159,122
; PRIOR FILING DATE: 1999-10-13
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 10
; LENGTH: 1902
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1902)
; US-10-086-464-10

```



```
QY 255 lLeuGlyPheSerIysSerThrPheThrTyrGluGluLeuAlaArgAlaThrAsnGlyPhe 275
Db 809 CCTCGGATTCACCAAGAGCACTTTTACTTACCAAGAGCTTCGGCTGCAACAGAGGGTT 868
QY 275 eSerGluAlaAsnLeuLeuGlyGlnGlyPheGlyTyrValHisIysGlyValLeuPr 295
Db 869 TAGCGATGCTAACCTTTTGGACAGGAGGATTTGGGTATGTCTCAAAAGAGTCTTGCC 928
QY 295 oSerGlyLeuGluValAlaValIysGlnLeuLeuValIglySerGlyGlnGlyGluArgG1 315
Db 929 TAGCGGAAAGAGTAGCAGTTAAGAGTTTAAAGCGGTAGCGGACAAGAGAGAGGGA 988
QY 315 uPheGluAlaGluValGluLeuIleSerArgValHisHisArgHisLeuValSerLeuVa 335
Db 989 GTTTCAAGCTGAGTGCATATCATATAGCCGTGTGCATCATCGGTATCTTGTCTTTGGT 1048
QY 335 lGlyTyrCysIleAlaGlyAlaIysArgLeuLeuValTyrGluPheValProAsnAsnAs 355
Db 1049 TGGATATTGATAGCTGATGACAGAGATGTGGTTTATGAGTTTGTCTTAACAAAC 1108
QY 355 nLeuGluLeuHisLeuHisGlyGlyArgProThrMetGluTrpSerThrArgLeuLy 375
Db 1109 TTTGGAATATCATCTTCATGGAAATCTTCCGGTAATGGAGTCTCCACTAGGTTCG 1168
QY 375 sIleAlaLeuGlySerAlaIysGlyLeuSerTyrLeuHisGluAspCysAsnProIysI1 395
Db 1169 TATCGCCTTAGGTGTCGAAAGGACTCGCTTACCTTCACGAAGACTGCCATCTCGGAT 1228
QY 395 eIleHisArgAspIleIysAlaSerAsnIleLeuIleAspPheIysPheGluAlaIysVa 415
Db 1229 CATTCACCGGACATCAAGTCTGCAATATCTCTTGACCTTCAACTTATGATCTATGGT 1288
QY 415 lAlaAspPheGlyLeuAlaIysIleAlaSerAspThrAsnThrHisValSerThrArgVa 435
Db 1289 GGCTGATTTTGGATTAGCTAAGTTAAACATCTGTATAACAACACTCATGTACTACTGTGT 1348
QY 435 lMetGlyThrPheGlyTyrLeuAlaProGluTyrAlaAlaSerGlyIysLeuThrGluLy 455
Db 1349 GATGGGAACCTTCGGATATCTAGCTCCAGAAATATGCTTCAAGCGGTAAATTAACCGAGAA 1408
QY 455 sSerAspValPheSerPheGlyValValLeuLeuGluLeuIleThrGlyArgArgProVa 475
Db 1409 ATCGATGTTTCTCTACCGAGTTATGTTATTTGGAACCTTATAACTGGAACACGCCGT 1468
QY 475 lAspAlaAsnAsnValTyrValAspAspSerLeuValAspTrpAlaArgProLeuLeuAs 495
Db 1469 TGAT--AATAGCATCACCATGGACGACACCTTAGTAGATTGGGCTCGGCTCTTATGGC 1525
QY 495 nArgAlaSerGluGlnGlyAspPheGluGlyLeuAlaAspAlaIysMetAsnAsnGlyTy 515
Db 1526 TCGCGCGCTAGAAGATGGAAACTTTAATAGACTCGCAGATCGCAGGGCTTGAAGCAACTA 1585
QY 515 rAspArgGluGluMetAlaArgMetValAlaCysAlaAlaCysValArgHisSerAl 535
Db 1586 CAACCCCAAGAAATGGCTCGAATGGTGACTTGTGCGCGCTGTAGCAATTCTGATTCGGG 1645
QY 535 aArgArgArgProArgMetSerGlnIleValArgAlaLeuGluGlyAsnValSerLeuSe 555
Db 1646 GCGTAAACGTCCTCAAGATAGCCAGATAGTAGAGCGGTAGAAGGAGAGTGTCTTAGA 1705
QY 555 rAspLeuAsnGluGlyMetArgProGlyGlnSerAsnValTyrSerSerTyrGlyGlySe 575
Db 1706 TGCTTTAAACCAAGAGTGTGAAGCCAGGACACAGTAACGTTTACGGGTCAVTTGGGAGCAAG 1765
QY 575 rThrAspTyrAspSerSerGlnTyrAsnGluAspMetLysIlePheArgGlyMetAlaLe 595
Db 1766 CTGGATTATAGTCAGACATCTTACAATGCAGACATGAAGAAATTCAGACAGATAGCTTT 1825
QY 595 uGlyThrGlnGluTyr-----AsnAlaThrGlyGluTyrSerAsnProThrSerAs 612
Db 1826 GTCGAGCCAAAGATTTCCAGCTAGCTGACTGTGAAGGAACATCTAGTAATGATCTAGAGA 1885
```

```
QY 612 pTyrGly 614
Db 1886 TATGGGA 1892
```

RESULT 9

```
US-10-425-114-9117
; Sequence 9117, Application US/10425114
; Publication No. US20040034886A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 9117
; LENGTH: 1424
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700833950_FLI
US-10-425-114-9117
```

```
Alignment Scores:
Pred. No.: 2,31e-104 Length: 1424
Score: 1737.50 Matches: 331
Percent Similarity: 88.73% Conservative: 47
Best Local Similarity: 77.70% Mismatches: 41
Query Match: 50.32% Indels: 7
DB: 13 Gaps: 4
```

US-10-086-464-2 (1-647) x US-10-425-114-9117 (1-1424)

```
QY 226 ProProProProProPheMetSerSerSerGly-----GlySerAspTyrSerAsp 243
Db 12 CCCCCACCCCGCTCTTTCATCAGCAGCAGTGGGATCTGGATCAACAAATTCAGGC 71
QY 244 ArgProValLeuProProProSerProGlyLeuValLeuGlyPheSerIysSerThrPhe 263
Db 72 GGTGAATTTCTCTCTCTCTCCAGGAATTCATTTGGGTTCTCTAAGAGCACATTC 131
QY 264 ThrTyrGluGluLeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuGlyGln 283
Db 132 ACGTATGAGGAGTTGGCAGCGCACTGATGGCTTCTCTGATGCCAACCTCTCTGGACAA 191
QY 284 GlyGlyPheGlyTyrValHisIysGlyValLeuProSerGlyIysGluValAlaValIys 303
Db 192 GGAGGATTTGGATATGTGCACAGAGGAATTTCTTCCCAACGCAAGGAGGTGGCAGTGAAG 251
QY 304 GlnLeuIysValGlySerGlyGlnGlyGluArgGluPheGlnAlaGluValGluIleIle 323
Db 252 CAATTGAAGGCTGGAAGCGGCAAGGGAGCGTGAATTCCAAGCTGAAGTTGAGATAATT 311
QY 324 SerArgValHisHisArgHisLeuValSerLeuValGlyTyrCysIleAlaGlyAlaIys 343
Db 312 AGCCGTGTCATCACAGCATCTGTGTTCTTTGGTGATCTGCACTCTGGTCCAG 371
QY 344 ArgLeuLeuValTyrGluPheValProAsnAsnLeuGluLeuHisLeuHisGlyGlu 363
Db 372 AGCTGCTTGTATGAATTTGTTCACCAACACACATTCGAATTCATTTGCATGGAAAA 431
QY 364 GlyArgProThrMetGluTrpSerThrArgLeuIysIleAlaLeuGlySerAlaIysGly 383
Db 432 GGAGAGACCTACCATGATTTGGCCCAAGACTAAGAAATTCCTTTAGGATCTGCTAAGGGA 491
QY 384 LeuSerTyrLeuHisGluAspCysAsnProLysIleIleHisArgAspIleLysAlaSer 403
```

```
Db 492 CTGGCGTATCTTCATGAAGATTGTCTATCTTAAGATCATCCATCGTGATATCAAAATCTGCC 551
Qy 404 AsnIleLeuIleAspPheLysPheGluAlaLysValAlaAspPheGlyLeuAlaLysIle 423
Db 552 AACATCTCTCGGATTTCAGTTTGAAGCAAGGTTGCAGATTTCCTTGGCTTCGCAAGTTT 611
Qy 424 AlaSerAspThrAsnThrHisValSerThrArgValMetGlyThrPheGlyTyrLeuAla 443
Db 612 TCTTCTGATGCAATACCCATGTTTCTACTCGAGTGATGGGACTTTTGGGTATTGGCT 671
Qy 444 ProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPheSerPheGlyVal 463
Db 672 CCAGAATATGCTCTAGTGGAAAACTCACAGAAATCAGATGTTTCTCTATGGGAATC 731
Qy 464 ValLeuLeuGluLeuIleThrGlyArgArgProValAspAlaAsnAsnValTyrValAsp 483
Db 732 ATGCTCTTGAGTTAATTAACCGGACGCGGCGAGTTGNTAAATCAAACTTACATGGAG 791
Qy 484 AspSerLeuValAspTrpAlaArgProLeuLeuLeuAsnArgAlaSerGluGlnGlyAspPhe 503
Db 792 GATAGTTTGGTAGATTGGCTTAGGCTTTGCTCACAGAGCTTTGGAAGAGGATGATTTT 851
Qy 504 GluGlyLeuAlaAspAlaLysMetAsnGlyTyrAspArgGluGluMetAlaArgMet 523
Db 852 GATTCTATTATGACCCCAAGGCTCCAGAAATGACTATGATCTCATGAGATGGCAGAAATG 911
Qy 524 ValAlaCysAlaAlaCysValArgHisSerAlaArgArgArgProArgMetSerGln 543
Db 912 GTGGCTTCTGCTGGGCTTGTATGCTATTCGTCATTCGGCAAGCGCTGACCAGATGAGCGAG 971
Qy 544 IleValArgAlaLeuGluGlyAsnValSerLeuSerAspLeuAsnGluGlyMetArgPro 563
Db 972 GTTGTCCGCGCTCTCGAAGGAGATGCTCTCTAGCAGATCTGAATGAAGGAATAAGACCT 1031
Qy 564 GlyGlnSerAsnValTyrSerSerTyrGlyGlySerThrAspTyrAspSerSerGlnTyr 583
Db 1032 GGACACAGCATATGATAGTTCTCAT---GAAAGCTCAGATTATGACACTGCACAGTAC 1088
Qy 584 AsnGluAspMetLysLysPheArgLysMetAlaLeuGlyThrGlnGluTyrAsnAlaThr 603
Db 1089 AAGGAAGACATGAAGAAAGTTGAGGAATGCGATTGGGAATCAGGAGTATGGTGCAAGC 1148
Qy 604 GlyGluTyrSerAsnProThrSerAspTyrGlyLeuTyrProSerGlySerSerSerGlu 623
Db 1149 AGTGAGTACAGTCCGCTACAGTGAGTATGTTTAAACCCATCAGGCTCAAGTAGTGAA 1208
Qy 624 Gly-----GlnThrThrArgGluMetGluMetGlyLysIleLysArgThrGlyGln 640
Db 1209 GCACAGAGCCGCCAAACCAAGGGAATGGAAATGGAAGATGAAG---AACAATCAA 1265
Qy 641 GlyTyrSerGlyProSer 646
Db 1266 GGTTTCAGTGAAGTTCT 1283

RESULT 10
US-10-086-464-13
; Sequence 13, Application US/10086464
; Publication No. US2002019218A1
; GENERAL INFORMATION:
; APPLICANT: GORING, Daphne R. et al.
; TITLE OF INVENTION: PROLINE-RICH EXTENSIN-LIKE RECEPTOR KINASES
; FILE REFERENCE: P 25,762-A USA
; CURRENT APPLICATION NUMBER: US/10/086,464
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 10/069,304
; PRIOR FILING DATE: 2002-02-19
; PRIOR APPLICATION NUMBER: PCT/CA00/00966
; PRIOR FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: US 60/149,466
; PRIOR FILING DATE: 1999-08-19
; PRIOR APPLICATION NUMBER: US 60/159,122
; PRIOR FILING DATE: 1999-10-13
; NUMBER OF SEQ ID NOS: 27
```

```
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 2025
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2025)
US-10-086-464-13

Alignment Scores:
Pred. No.: 6,22e-101 Length: 2025
Score: 1687.00 Matches: 368
Percent Similarity: 63.99% Conservative: 87
Best Local Similarity: 51.76% Mismatches: 142
Query Match: 48.86% Indels: 114
Dbs: 25

US-10-086-464-2 (1-647) x US-10-086-464-13 (1-2025)
Qy 1 MetSerAlaProSerProGlyThrGlySerProProSerProSerAsnSerThr 20
Db 16 GTGGATTTCATCTCTGCGCCCTGAAACC---TCAATGGGACACCCGTCNAACGGACCA 72
Qy 21 ThrThrThrProProAlaSerAlaProProProThrThrProSerSerProPro 40
Db 73 TCG-----CGTCTAATGAGTCATCGCGGCCAACACACCTTCTTCAACCAACCA 123
Qy 41 ProSerThrIleProThrSerProProSerSerArgSerThrProSerAlaProPro 60
Db 124 TCATCA-----ATATCTGCTCTCCGCCAGATATCTCCGCTCTTTTTCACCGCCG 174
Qy 61 ProSerProProThr-----ProSerThrProGlySerProProProLeu 75
Db 175 CTGCACCAACACGCAAGAAACGTCACCTCTACATCTCGTCTCTCATCGCCGCTGTT 234
Qy 76 -----ProGln-----ProSerProAlaPro-----Thr 84
Db 235 GTAGCTAATCCGTCACCGCAGACTCCACAGAAATCTTCTCCACCTGCACCTGAAGGCTCA 294
Qy 85 ThrProGlySerProProAla---ProValThrPro-----ProThrArgAsn 99
Db 295 ACTCTGTAAACGCGCCCTGCACCAACCAACCAACCGTCGAACCAATCACCAGGAAGACCA 354
Qy 100 ProProPro-----Pro-----Pro-----Pro----- 102
Db 355 ACTCTCTCTCTCTCTGTGTCATGATGACCGGAACAGAACCAATGCGGGAACCAACAC 414
Qy 103 -----SerValProGlyProProSerAsnProSerArg-----GluGlyGly 116
Db 415 AGAGACGGCTCCACACCATCACCGCTGTCAGGGAACAGAACTTCCGGTGACGGTGGC 474
Qy 117 SerProArgProProSerSerProProSerProProSerAspGlyLeuSerThr 136
Db 475 TCACCTTTCACCACTTCGTCGTAAGCCCTCTCAGAAATAGTGAGATTCAGACTCATCA 534
Qy 137 GlyValValGlyIleAlaIleGlyGlyValAlaLeuValIleValThrLeuIle 156
Db 535 TCG-----GGGCTT-----TTGCTTCTACTTGCAGTGTGATT 567
Qy 157 CysLeuLeuCysLysLysLysArgArgArgAspGluAspAla-----Tyr 172
Db 568 TGCATCTGTTGCAACAGGAAGAAAGAAATCTCTCAGGTCAACACCATGCACCTAC 627
Qy 173 TyrValProProProProProProGlyProLysAla---GlyGlyProTyrGlyGln 191
Db 628 TACAATAACAATCCTTATGAGGAGGACCCCTCAGTAATGCTGTGTTATTACAGGGA--- 684
Qy 192 GlnGlnGlnTrpArgGlnGlnAsnAlaThrProSerSerAspHisValValThrSerLeu 211
Db 685 -----ACACCTCAAGATCATGTGGTG---AATATG 711
Qy 212 ProProProProLysAlaProSerProProArgGlnProProProProProPro 231
```


US-10-424-599-57846

Alignment Scores:

Pred. No.: 3.37e-91 Length: 2655
Score: 1539.00 Matches: 341
Percent Similarity: 55.19% Conservative: 95
Best Local Similarity: 43.16% Mismatches: 141
Query Match: 44.57% Indels: 213
DB: 13 Gaps: 19

US-10-086-464-2 (1-647) x US-10-424-599-57846 (1-2655)

```
Qy 2 SerSerAlaProSerProGlyThrGlySerProProSerProSerAsnSer----- 19
Db 98 TCATCTTCTCCAAACA---AACACTTCCACCCACCATCACCTCCAGTCTCTCAGCCT 154
Qy 20 -----ThrThrThrThrProProAlaSerAlaPro---ProProThrThr 34
Db 155 AATCAAAACACAAACACAAACAAACCCCAATCCCTGCTCTCATCTCTCTCTGAC 214
Qy 35 ProSerSerProPro-----ProProSerThrThrProProSerProProSerSer 52
Db 215 CTTCTGCTCACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 274
Qy 53 ArgSerThrPro-----SerAlaProProProSerProProThrProSerThrPro 69
Db 275 TCACGTCTCTCTCCACCATCATCTACACCAACCACTTTGCCCTCTCTCTCTCTCT 334
Qy 70 GlySerProProProLeuPro-----GlnPro 78
Db 335 GATCTCTCACCAACCATTCACCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 394
Qy 79 SerProProAlaProThrProGlySerProGlySerProAlaProValThrProThrArg 98
Db 395 TCACCACCAAGTACAAACATCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 454
Qy 99 AsnProProProSerValPro-----GlyProProSerAsnProSerArgGluGly 116
Db 455 GAACTCCACCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 505
Qy 117 SerProArgPro-----ProSerSerPro 124
Db 506 TCCCTCTCTCTCAATCACTCTCCACACAAATTCACCTCTCTCTCTCTCTCTCTCT 565
Qy 125 SerProProSerProSer----- 130
Db 566 TCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 625
Qy 130 ----- 130
Db 626 GAAATCCACCAAAACTCTCTCTCATGATCTCTCTCTCTCTCTCTCTCTCTCTCT 685
Qy 130 ----- 130
Db 686 CCTAAACCTCTCTCTCTGATGTTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCT 745
Qy 130 ----- 130
Db 746 GACCTTTCAGGATCTTCACTCTCAGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 805
Qy 130 ----- 130
Db 806 GTGGGGGTCCGAAGGTGCTGCTACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 865
Qy 131 -----SerAspGlyLeu 134
Db 866 ACTAATGATGTTACTTAACAGTATGTTCTTCAACACACACACCTTACATCTCTG 925
Qy 135 SerThrGly-----ValValValGlyThrAlaValGlyValAlaLeuValLeu 152
Db 926 AGCACTGAGGATCTGCTGCTATTTGGAATTTGATGTTGTTTATTTCTCTCTCT 985
Qy 153 ValThrLeuIleCysLeuLeuCysAlaValGlyValGlyValGlyValGlyValGly 172
```

```
Db 986 GTTATGCTGTGCTGTTTGCACAGAAAAGAGAGGGAACAGGATCAGAGGTAGT 1045
Qy 173 TyrValProProProProProGlyProGlyProGlyProGlyProGlyProGlyPro 192
Db 1046 TATGCTGCTCTCTCTCTCCA----- 1063
Qy 193 GlnGlnTrpArgGlnGlnAsnAlaThrProProSerAspHisValValThrSerLeuPro 212
Db 1064 -----TTTACCTCATCCACCAATTCAGGTACCTTATTC 1096
Qy 213 ProProProLysAlaProSerProProArgGlnProProProProProProProPhe 232
Db 1097 TTGAGGCGCAGTCTCCGGC-----AACTTT 1123
Qy 233 MetSerSerSerGlySerAspTyrSerAspArgProValLeuProProProSerPro 252
Db 1124 TTAGGTAGTGGCTCTGGTGTGATTTT-----GTATATTCTCCATCAGAGCCT 1171
Qy 253 GlyLeuValLeuGlyPheSerLysSerThrPheThrTyrGluGluLeuAlaArgAlaThr 272
Db 1172 GGTGGTGTGTA---AGTAGTTCAAGATCATGTTGCATATGAAGAACTATTTCAGGTACA 1288
Qy 273 AsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyGlyPheGlyTyrValHisLysGly 292
Db 1229 AATGGGTTTTTCAGCACAAAATTTGTGGAGAGGTGGATTTGGCTGTGTTTATAAGGT 1288
Qy 293 ValLeuProSerGlyLysGluValAlaValLysGlnLeuLysValGlySerGlyGlnGly 312
Db 1289 TTGCTGATAGATGAGAGAGAGTAGCTGTGAAACAGCTCAAAATTTGGTGGTGGCAAGG 1348
Qy 313 GluArgGluPheGlnAlaGluValGluIleIleSerArgValHisHisArgHisLeuVal 332
Db 1349 GAACGCAATTCAGGCGCAAGTTGAGATTATTAGCCGTGTACATCATCGTCATCGTGT 1408
Qy 333 SerLeuValGlyTyrCysIleAlaGlyAlaValArgLeuLeuValTyrGluPheValPro 352
Db 1409 TCTTTAGTTGGTACTGTATATCCGACATCAGAGATTGCTGTATATGACTATGTTCCC 1468
Qy 353 AsnAsnAsnLeuGluLeuHisLeuHisGlyGluGlyArgProThrMetGluTrpSerThr 372
Db 1469 AACGATACTCTTATTACCATCTCCACGGTGAAATAGACCAGTTCTTAGATTGGCTACC 1528
Qy 373 ArgLeuLysIleAlaLeuGlySerAlaLysGlyLeuSerTyrLeuHisGluAspCysAsn 392
Db 1529 AGAGTCAAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1588
Qy 393 ProLysIleIleHisArgAspIleLysAlaSerAsnIleLeuIleAspPheLysPheGlu 412
Db 1589 CCAGCATTTATTATCAGATATTATTAAGTATCAACAACATCTTCTGATCTCAACTATGAA 1648
Qy 413 AlalysValAlaAspPheGlyLeuAlaLysIleAlaSerAspThrAsnThrHisValSer 432
Db 1649 GCTCAAGTTTCGGACTTTGGCTTTGCAAAATTTGGCATTAGATTCAATACATGTAAT 1708
Qy 433 ThrArgValMetGlyThrPheGlyTyrLeuAlaProGluTyrAlaAlaSerGlyLysLeu 452
Db 1709 ACAGTGTAAATGGGAACCTTTGGGTATATGGCAACCAATATGCAAGTGGAAAACTT 1768
Qy 453 ThrGluLysSerAspValPheSerPheGlyValValLeuLeuGluLeuThrGlyArg 472
Db 1769 ACTGAAAAGTCTGATGATATTCTTTTGGGGTGTGCTTTTGGAGCTAATTTACAGGTGG 1828
Qy 473 ArgProValAlaAspAlaAsnAsnValTyrValAspAspSerLeuValAspTrpAlaArgPro 492
Db 1829 AAGCCTGTAGATGCTCTCAACCAATTTGGTGTATGAGAGCTGGTTGAATGGGCTCGACCT 1888
Qy 493 LeuLeuAsnArgAlaSerGluGlnGlyAspPheGlyLeuAlaAspAlaLysMetAsn 512
Db 1889 CTGTTGAGGAGACCTTGAATGAGGACTTTGAAATTTTGGTGGATCCAGAGCTCGGG 1948
Qy 513 AsnGlyTyrAspArgGluGluMetAlaArgMetValAlaCysAlaAlaCysValArg 532
```


Db 1949 AAGAACTACGATAGAAATGTTTCGGATGATCGAGGTCGCCGACCTGTGTACGC 2008
Qy 533 HisSerAlaArgArgMetSerGlnIleValArgAlaLeuGluGlyAsnVal 552
Db 2009 CACTCATCGGTGAAGAGACCGCATGATGATGAGTGTGAGAGCTTTGATTCCTTGGAT 2068
Qy 553 SerLeuSerAspLeuAsnGluGlyMetArgProGlyGlnSerAsnValTyrSerSerTyr 572
Db 2069 GAGTTTACGGATCTCAATTAACGGAATGAACCGGACAGAGTTCGGTG----- 2116
Qy 573 GlyGlySerThrAspTyrAspSerSerGlnTyrAsnGluAspMetLysLysPheArgLys 592
Db 2117 -----TTTGATTCGGCGCAGCAATCTGCACAAATCAGAAATGTTTAGGAGG 2161
Qy 593 MetAlaLeuGlyThrGlnGluTyrAsnAlaThrGlyGlnTyrSerAsnProThrSerAsp 612
Db 2162 ATGGCTTTTGGAGGCCAGAT-----AGTTCCGGTTTCTTCATGAGTCTCAGAGTAGC 2215
Qy 613 Tyr-GlyLeuTyrProSerGlySerSerGluGlyGlnThrThrArgGluMetGluMe 632
Db 2216 TGGAGGAGTA-----GAGATCAGCAGCCCAACAACACTGTCTTC 2251
Qy 632 tGlyIysIleLysargThrGlyGlnGly 641
Db 2252 TCCAAAATAAAA---CTGGGCATGGGA 2276

RESULT 13

US-10-425-114-7339
; Sequence 7339, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 7339
; LENGTH: 2655
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700652889_FLI
US-10-425-114-7339

Alignment Scores:
Pred. No.: 3,37e-91 Length: 2655
Score: 1539.00 Matches: 341
Percent Similarity: 55.19% Conservative: 95
Best Local Similarity: 43.16% Mismatches: 141
Query Match: 44.57% Indels: 213
DB: 13 Gaps: 19

US-10-086-464-2 (1-647) x US-10-425-114-7339 (1-2655)

Qy 2 SerSerAlaProSerProGlyThrGlySerProProSerProProSerAsnSer----- 19
Db 98 TCATCTTCTCCACA---AACACTCCACCACATCACTCCAGTTCTTCTCAGCCT 154
Qy 20 -----ThrThrThrThrProProAlaSerAlaPro---ProProThrThr 34
Db 155 AATCAACACAAACACAAACACAAACCAATTCCTCTCTCTTCTCTCTGAC 214
Qy 35 ProSerSerProPro-----ProProSerThrIleProThrProProProSerSer 52
Db 215 CTTTCTGCTCCACCT 274

Qy 53 ArgSerThrPro-----SerAlaProProProSerProProThrProSerThrPro 69
Db 275 TCACTGTCTCTCCACCATCATCTACACCAACCAACTTTGCCCTCCATCACCACCT 334
Qy 70 GlySerProProProLeuPro-----GlnPro 78
Db 335 GATTCTCCACCATTTGCCACCTGCACTCTCAACCCCAACGTTTACAACATCTCCCCCC 394
Qy 79 SerProProAlaProThrThrProGlySerProAlaProValThrProThrArg 98
Db 395 TCACCACAGTTACAACATCTCCCTCTCAACCACTACCACTACCACTCTCTCTCCAGCT 454
Qy 99 AsnProProProSerValPro-----GlyProProSerAsnProSerArgGluGly 116
Db 455 GAAACTCCACTTCCCTTCCAAATTTGTCTCCACCATCTCCACCAGCC-----GGT 505
Qy 117 SerProArgPro-----ProSerSerPro 124
Db 506 TCCCTCTCTCTCAATCACCTCCCAACAATTCCACCTCTCTCTCTCGACCGATTTCCCA 565
Qy 125 SerProProSerProSer----- 130
Db 566 TCTCTCTCTCTCGGCCAATTTTCCAGGCCCAACCACTAGAACCTCTCCAGAGAAA 625
Qy 130 ----- 130
Db 626 GAAATCCACAAAACACTACTCTTCAATCATCATCTCTCTCTCTCTCTCTCTCTCTCT 685
Qy 130 ----- 130
Db 686 CTTAAACCTCTCTCTCTGATGTTCCCTCTCCATCCACATTTGCCCTTCAACTCTCTCTCA 745
Qy 130 ----- 130
Db 746 GACCTTCAGATCTTCACTCCAGCTTCTTGGCTGATCCCACTAATAAACAAGTG 805
Qy 130 ----- 130
Db 806 GTGGGGGTCCGAAGGTGTCGCTACCTCTCTTCCAACTGAGAAACCACTCTCTCTCTCTCT 865
Qy 131 -----SerAspGlyLeu 134
Db 866 ACTAATGATGGTACTAACAGTATGTCTTCAAAACAACACACTTTCACATTTCTGGAGGTG 925
Qy 135 SerThrGly-----ValValValGlyIleAlaIleGlyGlyValAlaLeuValIle 152
Db 926 AGCACTGGAGGATCTGGCTATTTGGAAATTTGATGTTTGTCTCTCTCTCTCTCTCTCT 985
Qy 153 ValThrLeuIleCysLeuLeuCysLysLysLysArgArgArgAspGluGluAspAlaTyr 172
Db 986 GTTATGGCTGTGTGTTTGCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1045
Qy 173 TyrValProProProProProProGlyProGlyProGlyProGlyProGlyGlnGln 192
Db 1046 TATGCTGCTCTCTCTCTCA----- 1063
Qy 193 GlnGlnTrpArgGlnGlnAsnAlaThrProProSerAspHisValValThrSerLeuPro 212
Db 1064 -----TTTACTCTATCCCACTTCCAGGTACCTTATTC 1096
Qy 213 ProProProLysAlaProSerProProArgGlnProProProProProProPhe 232
Db 1097 TTGAGGCCGAGTCTCGGCC-----AACTTT 1123
Qy 233 MetSerSerSerGlyGlySerAspTyrSerAspArgProValLeuProProProSerPro 252
Db 1124 TTAGTAGTGGCTCTGGTAGTATTT-----GTATATTCTCCATCAGAGCCT 1171
Qy 253 GlyLeuValLeuGlyPheSerLysSerThrPheThrTyrGluGluLeuAlaArgAlaThr 272
Db 1172 GGTGGTGTA---AGTAGTTCAAGATCATGTTTCAATATTAAGAACTATTTCAGCTACA 1288
Qy 273 AsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyPheGlyTyrValHisLysGly 292


```
Qy 243 AspArgProVal---LeuProProSerProGlyLeuValLeuGlyPheSerLysSer 261
Db 143 ACGGACCGCGCCGTTGCCCTCCCTCCCTACCAAACTTGTCTGGGCTCAAAAGGGGC 202
Qy 262 ThrPheThrTyrGluGluLeuAlaArgAlaThrPheGlyPheSerGluAlaAsnLeuLeu 281
Db 203 ACCTTCACTTACGAGGAATTAGCAGCTCCCAACCGAATTCATGATGCAAAATTTGATA 262
Qy 282 GlyGlnGlyGlyPheGlyTyrValHisLysGlyValLeuProSerGlyLysGluValAla 301
Db 263 GGACAAGTGTATTCGGCTATGTCCATAAGGGTGTGTGCTAGTGGAAAGAGTGGCA 322
Qy 302 ValLysGlnLeuLysValGlySerGlyGlnGlyGluArgGluPheGlnAlaGluValGlu 321
Db 323 GTTAAGAGCTTTAAAGCAGGTAGTGGCCAGAGAGCGAGAGTTCCAAGCTGAGATTGAC 382
Qy 322 IleIleSerArgValHisArgHisLeuValSerLeuValGlyTyrCysIleAlaGly 341
Db 383 ATCATTAGCCCGCTCCATCATCGCCATCTCTGTGCTACTTGTGGATATCAATTTCTGGT 442
Qy 342 AlAllysArgLeuLeuValTyrGluPheValProAsnAsnLeuGluLeuHisLeuHis 361
Db 443 GGCCAGAGAGTGTGGTCTATGATTTATTTCCCAATACACATTTGGATATCACCTCCAT 502
Qy 362 GlyGluGlyArgProThrMetGluTrpSerThrArgLeuLysIleAlaLeuGlySerAla 381
Db 503 GGAAGGGTAGGCCCTACCATGGATGGCCAACTAGATGCGGATTCGAATAGGATCCGCT 562
Qy 382 LysGlyLeuSerTyrLeuHisGluAspCysAsnProLysIleIleHisArgAspIleLys 401
Db 563 AAAGGGCTTGTTATCTTCATGAGAGCTGTCTCATCTCGTATTATTCATTCGCGATCAAA 622
Qy 402 AlaSerAsnIleLeuIleAspPheLysPheGluAlaLysValAlaAspPheGlyLeuAla 421
Db 623 CTGCAATATGCTTATGATGATAGCTTCGAAGCAAGGTGTGCTATTTGGATGGCT 682
Qy 422 LysIleAlaSerAspThrAsnThrHisValSerThrArgValMetGlyThrPheGlyTyr 441
Db 683 AAGTTGACTACTGATAATAATACTCATGTATCGACTCGTGTCTGCGGAACTTTCGGGTAC 742
Qy 442 LeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPheSerPhe 461
Db 743 CTAGCCCTCGAATATGATCAAGTGGAAAAATGACAGAGAAGTCTGACGTTTCTCATTT 802
Qy 462 GlyValValLeuLeuGluLeuIleThrGlyArgArgProValAspAlaAsnAsnValTyr 481
Db 803 GGAGTCATGCTATTGGAACTCATTAACCTGGGAAGCGACCTGTTGATCACACAATGCC--- 859
Qy 482 ValAspAspSerLeuValAspTrpAlaArgProLeuLeuAsnArgAla---SerGluGln 500
Db 860 ATGACGACGACAGCTAGTAGACTGGGCTCGACCACTTCTGACTCGTGGACTAGAGAGGAT 919
Qy 501 GlyAspPheGluGlyLeuAlaAspAlaLysMetAsnGlyTyrAspArgGluGluMet 520
Db 920 GGCAACTTTGAGAGTGTGGTGTGATTCATTTAGAGGGAACCTACGATCTCAAGAACTG 979
Qy 521 AlaArgMetValAlaCysAlaAlaCysValArgHisSerAlaArgArgArgProArg 540
Db 980 TCAAGATGGCAGCTTGTGCTCGCGGTAGCATTCGTCTATTCGCCAAAAAACGTCACAAA 1039
Qy 541 MetSerGlnIleValArgAlaLeuGluGlyAsnValSerLeuSerAspLeuAsnGluGly 560
Db 1040 ATGAGCCAGATTGTAAGAATATTGAGAGGAGATGTCTCACTGGATGACTTGAAGATGGG 1099
Qy 561 MetArgProGlyGlnSerAsnValTyrSerSerTyrGlyGlySerThrAspTyrAspSer 580
Db 1100 ATTAAGCCGGGCAAAATGTTGCTTTACAACTCTTCATCTAGCTCAGATCAGTATGACACA 1159
Qy 581 SerGlnTyrAsnGluAspMetLysLysPheArgLysMetAlaLeuGlyThrGlnGluTyr 600
Db 1160 ATGCAATACAAATGCTGATATGCAAGAGTTTCAGAAAGCGCAGTGTTCCTTAATAGCGGAA 1219
```

```
Qy 601 AsnAlaThrGlyGluTyrSerAsnProThrSerAspTyrGlyLeuTyrProSerGlySer 620
Db 1220 TTTGGCACCAGC-----AGTGGCTCC 1240
Qy 621 SerSerGlu 623
Db 1241 AGTGGTGA 1249
```

Search completed: May 13, 2004, 03:06:24
Job time : 728 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 12, 2004, 14:53:24 ; Search time 7358 Seconds
(without alignments)
11451.326 Million cell updates/sec

Title: US-10-086-464-1
Perfect score: 1944
Sequence: 1 atgtcctcggcgccgtctcc.....atagtggacctctctcttaa 1944

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 3470272 seqs, 21671516995 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rod.*
- 36: em_htg_mam.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1944	100.0	1944	6	AX088876	AX088876 Sequence
2	1944	100.0	1944	6	AX825703	AX825703 Sequence
3	1944	100.0	2189	6	AX825705	AX825705 Sequence
4	1944	100.0	2189	8	AF028699	AX028699 Brassica
5	66	3.4	1959	8	AX825738	AX825738 Sequence
6	66	3.4	1959	8	BT008400	BT008400 Arabidops
7	66	3.4	1959	8	BT008409	BT008409 Arabidops
8	66	3.4	2098	8	AY128792	AY128792 Arabidops
9	66	3.4	2116	8	AY056788	AY056788 Arabidops
10	66	3.4	2188	8	AY059901	AY059901 Arabidops
11	66	3.4	2190	8	AY093065	AY093065 Arabidops
12	66	3.4	2257	8	AF370509	AF370509 Arabidops
13	51	2.6	79706	8	AB020746	AB020746 Arabidops
14	47	2.4	2324	8	AY089024	AY089024 Arabidops
15	29	1.5	80393	8	AP000382	AP000382 Arabidops
16	26	1.3	1515	6	AX825736	AX825736 Sequence
17	26	1.3	47428	6	AX059540	AX059540 Sequence
18	26	1.3	107826	2	AC124968	AC124968 Medicago
19	26	1.3	145453	8	AC012477	AC012477 Genomic S
20	26	1.3	183147	8	AC012392	AC012392 Genomic S
21	26	1.3	199987	8	ATCHRIV15	AT161503 Arabidops
22	24	1.2	898	8	AK109201	AK109201 Oryza sat
23	24	1.2	1389	6	AX654306	AX654306 Sequence
24	24	1.2	1922	8	AK070323	AK070323 Oryza sat
25	24	1.2	88450	8	AP004622	AP004622 Oryza sat
26	23	1.2	256	11	GI2716	GI2716 sWSS2029 Eri
27	23	1.2	450	4	AF213398	AF213398 Ovis arie
28	23	1.2	897	6	AX546183	AX546183 Sequence
29	23	1.2	897	6	AX546184	AX546184 Sequence
30	23	1.2	897	6	AX546187	AX546187 Sequence
31	23	1.2	1185	8	AY224526	AY224526 Oryza sat
32	23	1.2	1281	6	AX546185	AX546185 Sequence
33	23	1.2	1620	9	HSU08818	U08818 Human activ
34	23	1.2	2016	9	HSU19348	U19348 Human (tpr-
35	23	1.2	3142	8	AY091180	AY091180 Arabidops
36	23	1.2	3407	8	AY064019	AY064019 Arabidops
37	23	1.2	3677	9	HSM808933	BX648782 Homo sapi
38	23	1.2	4419	4	AB118945	AB118945 Canis fam
39	23	1.2	4586	9	HSMETPRO	XS4559 Homo sapien
40	23	1.2	4620	6	AX546174	AX546174 Sequence
41	23	1.2	4626	6	I96185	I96185 Sequence 22
42	23	1.2	4626	9	HUMMETPOA	J02958 Human MET p
43	23	1.2	4626	11	GI8239	GI8239 sWSS842 Eri
44	23	1.2	4823	4	AB112434	AB112434 Bos tauru
45	23	1.2	32173	9	AC004416	AC004416 Homo sapi

ALIGNMENTS

RESULT 1	AX088876	AX088876	1944 bp	DNA	linear	PAT 17-MAR-2001
LOCUS	Sequence 1 from Patent WO0114563.					
DEFINITION	AX088876					
ACCESSION	AX088876.1	GI:13397639				
VERSION	Brassica napus (rape)					
KEYWORDS	Brassica napus					
SOURCE	Brassica napus					
ORGANISM	Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta;					
	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;					
	rosids; eurosids II; Brassicales; Brassicaceae; Brassica.					
REFERENCE	1					
AUTHORS	Goring,D. and Silva,N.					
TITLE	Proline-rich extensin-like receptor kinases					

Db	121	CGATCCACTATTCCGACATCTCTCTCTCTTCTTCGCTCTACACCTTCTGCTCTCTCT	180
Qy	181	CCATCTCCACCAACTCCATCTACGCGGGGATCTCCACCTCCTCTTCTCTCAGCGGCTCTCCA	240
Db	181	CGATCTCCACCAACTCCATCTACGCGGGGATCTCCACCTCCTCTTCTCTCAGCGGCTCTCCA	240
Qy	241	CCCGCTCCAACTACGCGCCGGATCTCCACCGCACCTGTTACTCTCTCTACTCGAAACCCCT	300
Db	241	CCCGCTCCAACTACGCGCCGGATCTCCACCGCACCTGTTACTCTCTCTACTCGAAACCCCT	300
Qy	301	CCACCTTCAGTCCAGGACCAACGTCCTTCTACGCGAGGAGGATCTCTCTCGACCT	360
Db	301	CCACCTTCAGTCCAGGACCAACGTCCTTCTACGCGAGGAGGATCTCTCTCGACCT	360
Qy	361	CCATCTTCTCCCTCGCGCGGCTCTCTTCTTCGACGGTTTATCAACAGGAGTGTTGGTG	420

Qy	421	GGAAATCGCCATCGGAGGAGTGCCTCTGCTTGATAGTACCTCTGATTTGTCTCCCTCTGT	480
Db	421	GGAAATCGCCATCGGAGGAGTGCCTCTGCTTGATAGTACCTCTGATTTGTCTCCCTCTGT	480
Qy	481	AAGAAGAAACGACGGAGAGACGAAGAAGATGCTTACTATGTTCTCTCGGCACCTCTCTCCT	540
Db	481	AAGAAGAAACGACGGAGAGACGAAGAAGATGCTTACTATGTTCTCTCGGCACCTCTCTCCT	540
Qy	541	GGTCCCAAAGCCGGAGGACCTTACGCTGGACAGCAGCAACAATGGCGGCACAAAAACGCA	600
Db	541	GGTCCCAAAGCCGGAGGACCTTACGCTGGACAGCAGCAACAATGGCGGCACAAAAACGCA	600
Qy	601	ACACCACCGCTCAGATCATGTGCTGACGTCACTACCAACCACCACTAAGGCTCCATCTCCA	660
Db	601	ACACCACCGCTCAGATCATGTGCTGACGTCACTACCAACCACCACTAAGGCTCCATCTCCA	660
Qy	661	CCACGGCAACCTCTCTCCACCTCCACCAACCGCTTTTCATGAGCAGACACGGCGGCTCCGAC	720
Db	661	CCACGGCAACCTCTCTCCACCTCCACCAACCGCTTTTCATGAGCAGACACGGCGGCTCCGAC	720

Db	721	TACTCGAGCGGTCCAGGTTCTTTCTCCACCGTCTCCAGGGCTGTGTAGGTTCTCCAAA	780
Qy	781	AGCACTTTCAATACAGGAGCTAGCTAGAGGCCAAATGGTTTCTCCGAGGCGAACTTG	840
Db	781	AGCACTTTCAATACAGGAGCTAGCTAGAGGCCAAATGGTTTCTCCGAGGCGAACTTG	840
Qy	841	TTAGGACAGGCGGGTTCGGTTACGTGTCACAAAGGTGTGTGCCTAGTGGGAAGAAGTT	900
Db	841	TTAGGACAGGCGGGTTCGGTTACGTGTCACAAAGGTGTGTGCCTAGTGGGAAGAAGTT	900
Qy	901	GCTGTCAAGCAGTTTGAAGTTTGGGAGTGTGTACGGGAGAGAGGAGGTTTCAAGGCAGAGTT	960
Db	901	GCTGTCAAGCAGTTTGAAGTTTGGGAGTGTGTACGGGAGAGAGGAGGTTTCAAGGCAGAGTT	960
Qy	961	GAGATCATCAGCAGAGTTTCAACACAGGCATCTGGTGTCCTCTGTGGTTATTCGATCGCC	1020
Db	961	GAGATCATCAGCAGAGTTTCAACACAGGCATCTGGTGTCCTCTGTGGTTATTCGATCGCC	1020

	GATGCTTCAACAGATTGCTTTATGATGTTTGTTCTTAACAACAATCTCGAGCTTCACCTC	1080
Qy	GGTGCCAAAAGATTGCTTTATGATGTTTGTTCTTAACAACAATCTCGAGCTTCACCTC	1080
Dd	GGTGCCAAAAGATTGCTTTATGATGTTTGTTCTTAACAACAATCTCGAGCTTCACCTC	1080
Qy	CATGGCGAGGACGGGCTTCAATGGAAATGGAGCACAGATTGAAGAATTGCTCTTGGATCT	1140
Dd	CATGGCGAGGACGGGCTTCAATGGAAATGGAGCACAGATTGAAGAATTGCTCTTGGATCT	1140
Qy	GCTAAAGGACTTTCTTATCTTCATGAAGATTGCAATCCTTAAATCATTCACCGTGATATC	1200
Dd	GCTAAAGGACTTTCTTATCTTCATGAAGATTGCAATCCTTAAATCATTCACCGTGATATC	1200
Qy	AAGGCTTCAAACATAATTGATAGATTTCAAGTTTGAAGCTAAGGTGCCTGATTTGGTCTT	1260
Dd	AAGGCTTCAAACATAATTGATAGATTTCAAGTTTGAAGCTAAGGTGCTGATTTGGTCTT	1260


```
Qy 1021 GGTGCCAAAAGATTGCTTGTCTATGAGTTTGTCTTAACAACAATCTCGAGCTTCACTC 1080
Db 1117 GGTGCCAAAAGATTGCTTGTCTATGAGTTTGTCTTAACAACAATCTCGAGCTTCACTC 1176
Qy 1081 CATGCGGAGGAGCGCTTACAAATGGAATGAGACACAGATTGAAGATTGCTTGGATCT 1140
Db 1177 CATGCGGAGGAGCGCTTACAAATGGAATGGAACACAGATTGAAGATTGCTTGGATCT 1236
Qy 1141 GCTAAGGACCTTCTTATCTTATCATGAAGATTGCAATCTTAAATCATTCACCGTGATATC 1200
Db 1237 GCTAAGGACCTTCTTATCTTATCATGAAGATTGCAATCTTAAATCATTCACCGTGATATC 1296
Qy 1201 AAGCTTCAAAATATGATAGATTCAAGTTTGAAGTAAAGTTGCTGATTTGGTCTT 1260
Db 1297 AAGCTTCAAAATATGATAGATTCAAGTTTGAAGTAAAGTTGCTGATTTGGTCTT 1356
Qy 1261 GCTAAGATTGCTTCTGATACAAAACAGCATATGATCAACACGTGTGATGGAACTTTGGG 1320
Db 1357 GCTAAGATTGCTTCTGATACAAAACAGCATATGATCAACACGTGTGATGGAACTTTGGG 1416
Qy 1321 TACTTGGCTCCGGAATACGCTGCAAGCGGAAGCTCACGGAAGCTCTGACGTTTCTCA 1380
Db 1417 TACTTGGCTCCGGAATACGCTGCAAGCGGAAGCTCACGGAAGCTCTGACGTTTCTCA 1476
Qy 1381 TTTGGCGTTGCTTTTGGAGCTCAATTAAGTCAAGCTGCAAGCGTTGATGCGCAACAATGTC 1440
Db 1477 TTTGGCGTTGCTTTTGGAGCTCAATTAAGTCAAGCTGCAAGCGTTGATGCGCAACAATGTC 1536
Qy 1441 TATGTAGATGACAGCTTATGATGATGCGGACGACCAATGCTTAAACCGAGCATCTGAGCAA 1500
Db 1537 TATGTAGATGACAGCTTATGATGATGCGGACGACCAATGCTTAAACCGAGCATCTGAGCAA 1596
Qy 1501 GGAGACTTTGAGGTTTATGCTGATGATCAAGATGATATGATGATGATGATGATGATGATG 1560
Db 1597 GGAGACTTTGAGGTTTATGCTGATGATCAAGATGATATGATGATGATGATGATGATGATG 1656
Qy 1561 GCTGCGATGTTGCTTGTGCTGCGCTTGTGTTGCCATTGCTGCGCGGAGACCTCGC 1620
Db 1657 GCTGCGATGTTGCTTGTGCTGCGCTTGTGTTGCCATTGCTGCGCGGAGACCTCGC 1716
Qy 1621 ATGAGCCAGATTGCTGCGGTTAGAGAAATGATATCACTGTCAGATCTTAAACGAAGGG 1680
Db 1717 ATGAGCCAGATTGCTGCGGTTAGAGAAATGATATCACTGTCAGATCTTAAACGAAGGG 1776
Qy 1681 ATGAGCCAGCTCAAGCAATGATATACAGCTCATACGAGGAGGACCGATTATGACTCG 1740
Db 1777 ATGAGCCAGCTCAAGCAATGATATACAGCTCATACGAGGAGGACCGATTATGACTCG 1836
Qy 1741 AGCCAGTACAATGAAGACATGAAGAGTTTATAGGAAAATGGCACTTGGAACTCAAGAGTAC 1800
Db 1837 AGCCAGTACAATGAAGACATGAAGAGTTTATAGGAAAATGGCACTTGGAACTCAAGAGTAC 1896
Qy 1801 AACGCCAGGTGAGTACAGTAATCCGACAGTATGATGACTGTACCCGTCTGGTTCA 1860
Db 1897 AACGCCAGGTGAGTACAGTAATCCGACAGTATGATGACTGTACCCGTCTGGTTCA 1956
Qy 1861 AGCAGCGAGGCGCAACACACCGCAATGAGATGGGAGATTTAGAGAACCGGTGAG 1920
Db 1957 AGCAGCGAGGCGCAACACACCGCAATGAGATGGGAGATTTAGAGAACCGGTGAG 2016
Qy 1921 GGTATAGTGACCTTCTCTTTAA 1944
Db 2017 GGTATAGTGACCTTCTCTTTAA 2040
```

RESULT 4

```
AY028699 LOCUS AY028699 2189 bp mRNA linear PLN 26-MAR-2001
DEFINITION Brassica napus receptor protein kinase PERK1 mRNA, complete cds.
ACCESSION AY028699
VERSION AY028699.1 GI:13447448
KEYWORDS Brassica napus (rape)
```

```
ORGANISM Brassica napus
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
TITLE roids; eurosids II; Brassicales; Brassicaceae; Brassica.
JOURNAL 1 (bases 1 to 2189)
AUTHORS Silva,N.F. and Goring,D.R.
TITLE Characterization of a novel Brassica napus receptor protein kinase
PERK1
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 2189)
TITLE Silva,N.F. and Goring,D.R.
JOURNAL Direct Submission
FEATURES Submitted (19-MAR-2001) Biology, York University, 4700 Keele
Location/Qualifiers Street, Toronto, Ontario M3J 1P3, Canada
source 1..2189
/mol_type="mRNA"
/db_xref="taxon:3708"
97..2040
/codon_start=1
/product="receptor protein kinase PERK1"
/protein_id="AAK21965.1"
/db_xref="GI:13447449"
/translaton="MSGAPSPGSGSPSPSPSNSTTTTTPPASAPPTTPSPSPSPSTT
PTSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP
SVSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP
KKRRRDEADAYVPPPPPPKAGPYGGQQQWQONQATPPSDHVVVTSLPPPKAP
SPRQPPPPPPPPHMSGSDSDRPLVPPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP
EANLGGGFGYVHKVLPSPKVEVAKLVGSGQGERFQAEVEIISRVRHRLVSL
VGYICAGAKLLVYEFVNNNLEHLHGEPTWEMSTRKIALGSAKGLSYLHEDCN
PKIIRDIKASNLIDPKPAKVADFLAKIASDNTNTHVTVMGTFGYLAPEVAASG
KLTKSPVFSFVVLLELTGRSPVDANNVYVDDSLVDMARPLANRASEQDPSGLAD
KXNNYDREMARWVACAAACVHRSHRRPRMSQIVRALEGNVSLDLNEGMPPGQS
KNVSSYGGSDYDSDQYNKRMALGTQBYNATGRYSNPTSDYGLYPSGSSESG
QTTREMEMGKIKRTGGYSGPFL"
ORIGIN
Query Match 100.0%; Score 1944; DB 8; Length 2189;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1944; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGTCTCTCGGCGCGCTCTCCGGGAGCTGGTTCGCTCCATCTCCACCATCAAACTCCACA 60
Db 97 ATGTCTCTCGGCGCGCTCTCCGGGAGCTGGTTCGCTCCATCTCCACCATCAAACTCCACA 156
Qy 61 ACCACCACTCTCTCCAGCTTCGCTCTCTCCACCACTCTCTCTCTCTCTCTCTCTCTCTCT 120
Db 157 ACCACCACTCTCTCCAGCTTCGCTCTCTCCACCACTCTCTCTCTCTCTCTCTCTCTCTCT 216
Qy 121 CCATCCACTATTCGACATCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180
Db 217 CCATCCACTATTCGACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 276
Qy 181 CCATCTCCACCACTCCATCTACCGCGGATCTCCACTCTCTCTCTCTCTCTCTCTCTCTCTCT 240
Db 277 CCATCTCCACCACTCCATCTACCGCGGATCTCCACTCTCTCTCTCTCTCTCTCTCTCTCTCT 336
Qy 241 CCGCTCCAACTACGCGCGGATCTCCACCGGACCTGTACTCTCTCTCTCTCTCTCTCTCTCTCT 300
Db 337 CCGCTCCAACTACGCGCGGATCTCCACCGGACCTGTACTCTCTCTCTCTCTCTCTCTCTCTCT 396
Qy 301 CCACCTTCAGTCCGAGGACCAACCGTCCAAATCTCTACGCGAAGGAGATCTCTCTCGACCT 360
Db 397 CCACCTTCAGTCCGAGGACCAACCGTCCAAATCTCTACGCGAAGGAGATCTCTCTCGACCT 456
Qy 361 CCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 420
Db 457 CCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 516
Qy 421 GGAATCGCCATCGGAGGAGTGGTCTGCTCTGTGTGATGACTCTGTATTTGCTCTCTCTCTGT 480
Db 517 GGAATCGCCATCGGAGGAGTGGTCTGCTCTGTGTGATGACTCTGTATTTGCTCTCTCTCTGT 576
```

```
QY 481 AAGAAGAAACGACGAGAGACGAGAGATGCTTACTATGTTCTCTCCGCCACCTCCTCCT 540
Db 577 AAGAAGAAACGACGAGAGACGAGAGATGCTTACTATGTTCTCTCCGCCACCTCCTCCT 636
QY 541 GGTCCAAAGCGGAGGACCTTACGGTGGACAGAGCAACAATGGCGCAACAAAACGCA 600
Db 637 GGTCCAAAGCGGAGGACCTTACGGTGGACAGAGCAACAATGGCGCAACAAAACGCA 696
QY 601 ACACCACCGTCAGATCATGTGCTGAGCGTCACTACCAACCAACCACTTAAGGCTCCATCTCCA 660
Db 697 ACACCACCGTCAGATCATGTGCTGAGCGTCACTACCAACCAACCACTTAAGGCTCCATCTCCA 756
QY 661 CCAACGGAACCTCTCCACCTCCACACCGGCTTTCATGAGCAGAGCGGCGGCTCCGAC 720
Db 757 CCAACGGAACCTCTCCACCTCCACACCGGCTTTCATGAGCAGAGCGGCGGCTCCGAC 816
QY 721 TACTCGGACCGTCCAGTTCTTCTCCACCGTCTCCAGGGCTTGTGTTAGGCTTCTCCAAA 780
Db 817 TACTCGGACCGTCCAGTTCTTCTCCACCGTCTCCAGGGCTTGTGTTAGGCTTCTCCAAA 876
QY 781 AGCACTTTTCAATACAGAGGAGTCTAGAGCCCAACCAATGGTTTCTCCGAGGCGAACTTG 840
Db 877 AGCACTTTTCAATACAGAGGAGTCTAGAGCCCAACCAATGGTTTCTCCGAGGCGAACTTG 936
QY 841 TTAGGACAAAGGGGTTCCGGTTACGTCACAAAGGTGTGTCCTAGTGGGAAAGAGTT 900
Db 937 TTAGGACAAAGGGGTTCCGGTTACGTCACAAAGGTGTGTCCTAGTGGGAAAGAGTT 996
QY 901 GCTGTGAAGCAGTTGAAAGTTGGGAGTGTGTCAGGAGAGAGGGATTTTCAGCAGAGGTT 960
Db 997 GCTGTGAAGCAGTTGAAAGTTGGGAGTGTGTCAGGAGAGAGGGATTTTCAGCAGAGGTT 1056
QY 961 GAGATCATCAGCAGAGTTTCCACACAGGCATCTGGTGTCTCTTGTGTTGTTATTCATCGCC 1020
Db 1057 GAGATCATCAGCAGAGTTTCCACACAGGCATCTGGTGTCTCTTGTGTTATTCATCGCC 1116
QY 1021 GGTGCCAAAAGATTGCTTGTCTATGAGTTGTTCTTAAACAATCTCGAGCTTCACCTC 1080
Db 1117 GGTGCCAAAAGATTGCTTGTCTATGAGTTGTTCTTAAACAATCTCGAGCTTCACCTC 1176
QY 1081 CATGCGAGGAGCGGCTTACAAATGGAATGGAGCACCAGATTGAAGATTGCTCTGGATCT 1140
Db 1177 CATGCGAGGAGCGGCTTACAAATGGAATGGAGCACCAGATTGAAGATTGCTCTGGATCT 1236
QY 1141 GCTAAAGGACCTTCTTATCTTTCATGAAGATTGCAATCTTAAATCAATTCACCGTGATATC 1200
Db 1237 GCTAAAGGACCTTCTTATCTTTCATGAAGATTGCAATCTTAAATCAATTCACCGTGATATC 1296
QY 1201 AAGGCTTCAACATATTGATGATTTCAAGTTTGAAGCTTAAGTTGCTGATTTGGTCTT 1260
Db 1297 AAGGCTTCAACATATTGATGATTTCAAGTTTGAAGCTTAAGTTGCTGATTTGGTCTT 1356
QY 1261 GCTAAGATTGCTTCTGATACAAACACGATGTATCAACACGTTGATGGAACTTTGGG 1320
Db 1357 GCTAAGATTGCTTCTGATACAAACACGATGTATCAACACGTTGATGGAACTTTGGG 1416
QY 1321 TACTTGGCTCCGGAATAGCGTCAAGCGGAAGGCTCACGGAGAAGTCTGACGTTTCTTCA 1380
Db 1417 TACTTGGCTCCGGAATAGCGTCAAGCGGAAGGCTCACGGAGAAGTCTGACGTTTCTTCA 1476
QY 1381 TTTGGCGTTGCTTTTGGAGCTCATTAATGGAAGTGTGATGCGCAACATGTC 1440
Db 1477 TTTGGCGTTGCTTTTGGAGCTCATTAATGGAAGTGTGATGCGCAACATGTC 1536
QY 1441 TATGTAGATGACGCTTATGTTGACCTGGGACGACCACTTCTTAAACGAGCATCTGAGCAA 1500
Db 1537 TATGTAGATGACGCTTATGTTGACCTGGGACGACCACTTCTTAAACGAGCATCTGAGCAA 1596
QY 1501 GGAGACTTTGAGGGTTTGGCTGATGCAAGAGATGAATAATGGGTATGACAGAGAGAGATG 1560
Db 1597 GGAGACTTTGAGGGTTTGGCTGATGCAAGAGATGAATAATGGGTATGACAGAGAGAGATG 1656
```

```
QY 1561 GCTCGCATGGTTGCTTGTGCTGGGCTTGTGTTCCGCAATTCAGCTCCGCGCAGACCTCGC 1620
Db 1657 GCTCGCATGGTTGCTTGTGCTGGGCTTGTGTTCCGCAATTCAGCTCCGCGCAGACCTCGC 1716
QY 1621 ATGAGCCAGATTGTCGTCGTCGTTAGAAAGAAATGTATCACTGTGTCAGATCTTTAAACGAAGG 1680
Db 1717 ATGAGCCAGATTGTCGTCGTCGTTAGAAAGAAATGTATCACTGTGTCAGATCTTTAAACGAAGG 1776
QY 1681 ATGAGCAGAGTCAAGCAATGTATACAGCTCATACGAGAGAGACCGATTATGACTCG 1740
Db 1777 ATGAGCAGAGTCAAGCAATGTATACAGCTCATACGAGAGAGACCGATTATGACTCG 1836
QY 1741 AGCCAGTACAATGAAGACATGAAGCTTTAGGAAAATGGCAGTTGGAACCTCAAGAGTAC 1800
Db 1837 AGCCAGTACAATGAAGACATGAAGCTTTAGGAAAATGGCAGTTGGAACCTCAAGAGTAC 1896
QY 1801 AACGCCACGGGTGAGTACAGTAATCCGACCAGTCACTATGAGCTGTATCCCGTCTGGTTCA 1860
Db 1897 AACGCCACGGGTGAGTACAGTAATCCGACCAGTCACTATGAGCTGTATCCCGTCTGGTTCA 1956
QY 1861 AGCAGCAGGCGCCAAACCAACACGCGAAATGGAGATGGGGAAGATTAAAGAGAACCGTTCAG 1920
Db 1957 AGCAGCAGGCGCCAAACCAACACGCGAAATGGAGATGGGGAAGATTAAAGAGAACCGTTCAG 2016
QY 1921 GGTATATAGTGACCTTCTCTTTAA 1944
Db 2017 GGTATATAGTGACCTTCTCTTTAA 2040

RESULT 5
AX825738
LOCUS AX825738
DEFINITION Sequence 36 from Patent WO03072763.
ACCESSION AX825738
VERSION AX825738.1 GI:39751255
KEYWORDS
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsiis.
REFERENCE
1
AUTHORS Goring, D., Silva, N. and Haffani, Y. Z.
TITLE Increasing plant seed production
JOURNAL Patent: WO 03072763-A 36 04-SEP-2003;
Goring, Daphne (CA) ; Silva, Nancy (CA) ; Haffani, Yosr, Z. (CA)
FEATURES
source
1. 1959
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"
ORIGIN
Query Match 3.4%; Score 66; DB 6; Length 1959;
Best Local Similarity 100.0%; Pred. No. 4.6e-24;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1432 AACAAATGCTATGTAGATGACGCTTAGTTGACTGGGACGACCACTTGTAAACCGAGCA 1491
Db 1447 AACAAATGCTATGTAGATGACGCTTAGTTGACTGGGACGACCACTTGTAAACCGAGCA 1506
QY 1492 TCTGAG 1497
Db 1507 TCTGAG 1512

RESULT 6
BT008400
LOCUS BT008400
DEFINITION Arabidopsis thaliana At3g24550 gene, complete cds.
ACCESSION BT008400
VERSION BT008400.1 GI:30725473
KEYWORDS FLI_CDNA.
```


TITLE
JOURNAL
REFERENCE
AUTHORS

Arabidopsis cDNA clones
 2 (bases 1 to 2116)
 Cheuk,R., Chen,H., Kim,C.J., Koesema,E., Meyers,M.C., Banh,J.,
 Bowser,L., Carninci,P., Dale,J.M., Goldsmith,A.D., Hayaishizaki,Y.,
 Ishida,J., Jiang,P.X., Jones,T., Kamiya,A., Karlin-Neumann,G.,
 Kawai,J., Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M.,
 Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Pham,P.K.,
 Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C.,
 Toriumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K.,
 Davis,R.W., Theologis,A. and Ecker,J.R.

TITLE
JOURNAL

Direct Submission
 Submitted (11-SEP-2001) Salk Institute Genomic Analysis Laboratory
 (SIGNAL), Plant Biology Laboratory, The Salk Institute for
 Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
 USA

COMMENT

RIKEN Genomic Sciences Center (GSC) members carried out the
 collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
 Arabidopsis Full-Length cDNA') : Seki,M., Narusaka,M., Ishida,J.,
 Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
 Hayaishizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGEC (SSP) Consortium members carried out the
 sequencing and annotation of the RAFL cDNAs: Cheuk,R., Chen,H.,
 Kim,C.J., Koesema,E., Meyers,M.C., Shinn,P., Banh,J., Bowser,L.,
 Dale,J.M., Goldsmith,A.D., Jiang,P.X., Jones,T., Karlin-Neumann,G.,
 Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M., Nguyen,M.,
 Onodera,C.S., Palm,C.J., Pham,P.K., Quach,H.L., Southwick,A.,
 Tang,C.C., Toriumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S.,
 Davis,R.W., Theologis,A., and Ecker,J.R.

FEATURES
 source
 1. .2116
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /chromosome="3"
 /clone="RAFL09-19-J24 (R14362)"
 /ecotype="Columbia"
 1. .36
 37. .1995
 /note="putative protein kinase"
 /codon_start=1
 /product="AT3g24550/MOB24_8"
 /protein_id="AAL10479.1"
 /db_xref="GI:15983765"
 /translation="MSTAPSGTTPSPSPPTNSTTTTPPAASPPPTTPSSPP
 PSESTNSTPSPSPPLPSPSPGSLTLPPLPQSPSPAPITPSPSPPTPNRSP
 PPNQGPNTPSGTPRTPTNTKPSPSDGLSTGVVGIAGVAILVILTLCL
 LCKKRRRHDDAAAYVPPPPSPGKAGPYGQQQYQQQNASRPSDNHVTSLPP
 PKPPPRKPPPPPPPPAFMSSGSDYSLPVLPPSPGLVLFKSTFTYELSLR
 TNGFSEANLLGGQGVVHGKILPSGKEVAVKOLKAGSGGGEREFOAEVRIISRVRH
 HLVSILGYCMAGVQRLVYFVFNPNLEFHLHGKGRPTMSTRLKALGAKGLSVL
 HEDCNPKIHRDKASNILIDFKPEAKVADFGIAKIASDTNTHVTRVMGTFGYLAPE
 YASGKTKSDVFSFVLELITRRPVDANNVYDDSLVDMWARLLNRAEEGDF
 EGLADSKMGNEYDREEMARMVACAAACVRSARRRPMQIVRALEGNVSLDNEGM
 RPSGHSNYSSGGSTDYDTSQYNDMDIKFRKMGALGTQEGYGTGTSNPTSDYGLYPSG
 SSEGQATREMEWGKIKKTGGYSGPSL"
 1996. .2116

3'UTR
ORIGIN

Query Match 3.4%; Score 66; DB 8; Length 2116;
 Best Local Similarity 100.0%; Pred. No. 4.6e-24;
 Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1432 AACATCTCTATGATGACAGCTTAGTTCAGTGGCAGCACCATTCCTTAACCGAGCA 1491
 |||||
 Db 1483 AACATCTCTATGATGACAGCTTAGTTCAGTGGCAGCACCATTCCTTAACCGAGCA 1542
 |||||

Qy 1492 TCTGAG 1497

Db 1543 TCTGAG 1548
 |||||

RESULT 10
AY059901
LOCUS
DEFINITION
 Arabidopsis thaliana protein kinase-like protein (MOB24.13) mRNA,
 complete cds.
ACCESSION
 AY059901
VERSION
 AY059901.1 GI:16649062
FLU CDNA.
SOURCE
ORGANISM
 Arabidopsis thaliana (thale cress)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopses.
 1 (bases 1 to 2188)
 Nguyen,M., Karlin-Neumann,G., Southwick,A., Lam,B., Miranda,M.,
 Palm,C.J., Bowser,L., Jones,T., Banh,J., Hayaishizaki,Y., Ishida,J., Kamiya,A.,
 Cheuk,R., Chung,M.K., Hayaishizaki,Y., Ishida,J., Kamiya,A.,
 Kawai,J., Kim,C., Lin,J., Liu,S.X., Narusaka,M., Pham,P.K.,
 Sakano,H., Sakurai,T., Satou,M., Seki,M., Shinn,P., Yamada,K.,
 Shinozaki,K., Ecker,J., Theologis,A. and Davis,R.W.
Direct Submission
 Submitted (22-OCT-2001) DNA Sequencing and Technology Center,
 Stanford University, 855 California Avenue, Palo Alto, CA 94304,
 USA
 e-mail for correspondence: arab@sequence.stanford.edu

TITLE
JOURNAL

COMMENT

RIKEN Genomic Sciences Center (GSC) members carried out the
 collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
 Arabidopsis Full-Length cDNA') : Seki,M., Narusaka,M., Ishida,J.,
 Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
 Hayaishizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGEC (SSP) Consortium members carried out the
 sequencing and annotation of the RAFL cDNAs: Nguyen,M.,
 Southwick,A., Karlin-Neumann,G., Lam,B., Miranda,M., Palm,C.J.,
 Bowser,L., Jones,T., Banh,J., Chen,H., Cheuk,R., Chung,M.K.,
 Kim,C., Lin,J., Liu,S.X., Pham,P.K., Sakano,H., Shinn,P.,
 Yamada,K., Ecker,J., Theologis,A. and Davis,R.W.

Nguyen,M. (SSP/Stanford) and Seki,M. (RIKEN GSC) contributed
 equally to this work. Shinozaki,K. (RIKEN GSC) and Davis,R.W.
 (SSP/Stanford) contributed equally to this work as PIs.

FEATURES
 source
 Location/Qualifiers
 1. .2188
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /chromosome="III"
 /clone="RAFL09-06-A17"
 /note="This clone is in pBluescript"
 1. .2188
 /gene="MOB24.13"
 32. .1990
 /genes="MOB24.13"
 /codon_start=1
 /product="protein kinase-like protein"
 /protein_id="AAL24383.1"
 /db_xref="GI:16649063"
 /translation="MSTAPSGTTPSPSPPTNSTTTTPPAASPPPTTPSSPP
 PSESTNSTPSPSPPLPSPSPGSLTLPPLPQSPSPAPITPSPSPPTPNRSP
 PPNQGPNTPSGTPRTPTNTKPSPSDGLSTGVVGIAGVAILVILTLCL
 LCKKRRRHDDAAAYVPPPPSPGKAGPYGQQQYQQQNASRPSDNHVTSLPP
 PKPPPRKPPPPPPAFMSSGSDYSLPVLPPSPGLVLFKSTFTYELSLR
 TNGFSEANLLGGQGVVHGKILPSGKEVAVKOLKAGSGGGEREFOAEVRIISRVRH
 HLVSILGYCMAGVQRLVYFVFNPNLEFHLHGKGRPTMSTRLKALGAKGLSVL
 HEDCNPKIHRDKASNILIDFKPEAKVADFGIAKIASDTNTHVTRVMGTFGYLAPE
 YASGKTKSDVFSFVLELITRRPVDANNVYDDSLVDMWARLLNRAEEGDF
 EGLADSKMGNEYDREEMARMVACAAACVRSARRRPMQIVRALEGNVSLDNEGM
 RPSGHSNYSSGGSTDYDTSQYNDMDIKFRKMGALGTQEGYGTGTSNPTSDYGLYPSG


```
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="III"
/clone="RAF11-07-N23"
/ecotype="Columbia"
/note="This clone is in pBluescript"
1..2257
/gene="MOB24.13"
107..2065
/gene="MOB24.13"
/codon_start=1
/product="protein kinase-like protein"
/protein_id="AAK3886.1"
/db_xref="GI:13877617"
/translation="MSTAPSPGTTTPTSPSPPTTSTTTTTPPAASSPPPTTPTSPSP
PSPSTNTPSPSPSLPPLPSPSPGSLTPPIQPSPSAPITPSPSPPTSPNPRP
PSPNQPPNTSGTPTPTSPNTKSPSPSDSLSTGVVGIAIGGVAILLILCL
LCKRRRRHDEAAYVYVPPPPSPGKAGPYGQQYQQQNASRPSDHHVTSLLP
PKPPSPRPPPPPPAFMSSSGSDYSDLPVLPSPSPGLVLGFSKSTYEISLRA
TNGFSEANLLQGQGVYHKGILPSGKEVAVKQKAGSGGQREFQAEVEIISRVHR
HLVSLIGVMAGVORLLVYFVPPNNLEFHLHGKGRPTMWSITLKIALGSAKGLYL
HEDCNPKLIHRDIKASNILIDFKEAKVADFLAKIASDITNTVSTVMTGTFGLAPE
YAAAGKLEKSDVFSGVVLELLITGRPVDDANNVYVDDSLVDWARPLNRAEEGDF
EGLADSKMGNEVDREEMARVACAAACVRHSARRPRMSQIVRALEGNVSLDINEGM
RPGHSNVSYSGSTDYDTQYNDMDIKFRKALGTQBYGTTGYSNPTSDYGLYPSG
SSSEGGATREMKGKIKKTGGYSGPSL"

FEATURES             source
CDS
1..79706
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia"
/db_xref="taxon:3702"
/chromosome="3"
/clone="MOB24"
/clone_lib="Mitsui P1"
2398..3399
/note="gene id:MOB24.1
unknown protein"
/evidence=not experimental
/protein_id="BAB01995.1"
/db_xref="GI:9294038"
/translation="MGDSDETCVPKKFYSLNQNHMPYSRPIPKRHAYVNEEDED
VEGEEDPGQYIRGNRFQKQKPNKVSQGFAGPSDAKVAYDRQEQAFVLLEVM
GDRFLQGRSLRNEDEVAEKVSEBELMEKSETQCRMIDDLKRYREKKIKVEKS
GLSSKWSFNKLDMLLCKPSKDLGACVDSGEFVNTKVLDSNDFDEKMDSP
GSESEDEDEVEYRKVNDAAYSKMLADSVRFKGYEKMEKSKKQMKLEKQNR
ADFORDLELQKQIVDRAQSEIARLEEEENHHGGDDDESEDEMDVNLSDS"
complement(join(3794..4276,4363..4490,4577..4673,
4757..5179,5354..5548,5693..5970,6088..6214,6297..6496,
6601..6648,6743..6776,6882..7013,7464..7559,7677..7743,
8162..8217,8290..8385,8470..8998,9122..9462))
/note="gene id:MOB24.2"
/codon_start=1
/evidence=not experimental
/product="DNA_repair protein MutS"
/protein_id="BAB01996.1"
/db_xref="GI:9294039"
/translation="MQQRSLTSFFQKTAATTKGLVSDAASGGSGGPRFNVKEG
DAKGDASVRFVSVSVDEVRGTDTPPEKVRPRLPSGFKPAESAGDASSLFSNMHKF
VKYDDRCGERSEEDVPLNDSLLCKANDVIQPSNNGKTQERNHAFSFGRAEL
RSVEDLGVDGVPGPTPGMRPRASRLKRVLEDMTFKEDKVPVDSNKKLMLQDPV
CGEKEVNGTGFKEWLESRIIDANRRRDPDLYDRKTLHIPDPVKFMSASQKQYS
VKSEYNDIVLFFKGFYELYELDABELKELDKMTGMSGVKQVQVIGISGIDEAV
KLLARGYKQSTVYGFADCAALRFVGSISDDASCAALGALLMQVSPKEVLYDSK
IKELMELKCVSTVYGFADCAALRFVGSISDDASCAALGALLMQVSPKEVLYDSK
GLSREQAALRKTYLTGTAVALAPVPQMGDTDAAGVNIIESNGYFKSSSSSNCA
VGLNCGDVALSGLYKLDNCVSTGKLLRNWICHKIDVESINRKLIDGGTWNLEIF
NNSCDGGPGSTGLYKLDNCVSTGKLLRNWICHKIDVESINRKLIDGGTWNLEIF
MIDTLLQKHLPLDLRLGRKISSVSSASVLPALGKVKLVKQVAFNGYQVQVTDDE
IDTLLQKHLPLDLRLGRKISSVSSASVLPALGKVKLVKQVAFNGYQVQVTDDE
NAETLLLELFIERATOWSEVHTTISCLDLVLSFAIAASLSAGSARVPVIFPESEAT
DQNKTKGILKIQGLWHFPAVADQQLFPVNDILLGEARRSGSIHPSRLLLTGPMN
GGKSTLLRATCLAVIFAQGCYVPCSEISLVDITIFTRUGASDILLGSLTFLVET
ETASVQNATQDSLVLDLBELGRTSTFDGVIAYSVFRLHVEKQVCMFLQVLMAGI
KEFASPRVTSKHMCAFKRSYQPRGCDQDLVFLYRLTEGACPSYGIQVAMAGI
PNOVETATGAAQMKRSIGENPKSELSEPSSESHEDMLKSLVIGISRVANNAPICE
DYDVTLCFLWHEITKSSYCVPK"
10131..10577
/note="gene id:MOB24.3"
/codon_start=1
/evidence=not experimental
/product="ethylene-responsive transcriptional
```


frame shifts in a coding region. A sequence is considered to be 5'-truncated if it lacks the translation initiation start (ATG). A sequence is considered to be 3'-truncated if it lacks the C-terminal end of the encoded protein. Please note that these cDNA sequences are derived from the Ws or Laer ecotypes and therefore may contain polymorphisms when compared to sequences from Col-0. Genset carried out the library production and sequencing of the full-length clones. Ceres, Inc. carried out the clustering of the 5' sequences, selection of clones, and sequence assembly.

FEATURES

Location/Qualifiers
1..2324
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="17909"

ORIGIN

Query Match 2.4%; Score 47; DB 8; Length 2324;
Best Local Similarity 100.0%; Pred. No. 1.4e-13;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1595 GCTGTGTTCCGACCTTACGCTCGCGCAGACCTCGCATGACCCAGAT 1631
|||||
Db 1700 GCTGTGTTCCGACCTTACGCTCGCGCAGACCTCGCATGACCCAGAT 1746
|||||

RESULT 15

AP000382/c 80393 bp DNA linear PLN 27-DEC-2000
LOCUS Arabidopsis thaliana genomic DNA, chromosome 3, TAC clone:K7M2.
DEFINITION AP000382 BA000014
ACCESSION AP000382.1 GI:5672513
VERSION

KEYWORDS

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM

Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

REFERENCE

AUTHORS Kaneko, T., Katoh, T., Sato, S., Nakamura, A., Asamizu, E. and Tabata, S.
TITLE Structural analysis of Arabidopsis thaliana chromosome 3. II.
Sequence features of the 4,251,695 bp regions covered by 90 P1, TAC
and BAC clones
JOURNAL DNA Res. 7 (3), 217-221 (2000)

MEDLINE

20363099

PUBMED

10907853

REFERENCE

2 (bases 1 to 80393)
Kaneko, T., Katoh, T., Sato, S., Nakamura, Y., Asamizu, E. and Tabata, S.

TITLE

Direct Submission
Submitted (28-JUL-1999) Yasukazu Nakamura, Karusa DNA Research

Institute, Department of Plant Gene Research; 1532-3, Yana,

Kisarazu, Chiba 292-0812, Japan (E-mail: ynakamu@kazusa.or.jp,

Tel: 81-438-52-3935, Fax: 81-438-52-3934)

Address for correspondence: kaos@kazusa.or.jp

For the latest information on annotation of this clone, please see

http://www.kazusa.or.jp/kaos/cgi-bin/aggd_graph.cgi?c=K7M2

Genes with similarity to proteins in the databases are described in

'product' or 'note' qualifiers. Genes that have no significant

protein similarity are described as 'unknown protein'.

The software programs used to predict genes include: Grail

(Informatics Group, Oak Ridge National Laboratory,

<http://compbio.ornl.gov/Grail-1.3/>),

GENSCAN (Chris Burge, MIT, <http://CCR-081.mit.edu/GENSCAN.html>),

NetGene2 (S.M. Hebsgaard, et al., CBS, Technical University of

Denmark, <http://www.cbs.dtu.dk/services/NetGene2/>) and

SplicePredictor (Volker Brendel, Stanford University,

<http://gremlin1.zool.iastate.edu/cgi-bin/sp.cgi>).

Genes encoding tRNAs are predicted by tRNAscan-SE

(Sean Eddy, Washington University School of Medicine, St. Louis,

<http://genome.wustl.edu/eddy/tRNAscan-SE/>).

This sequence may not be the entire insert of this clone. It may be

shorter because we remove overlaps between neighboring submissions.

The 5' clone is K13K6 and the 3' clone is MXPs.

FEATURES

source

Location/Qualifiers

1..80393
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia"
/db_xref="taxon:3702"
/chromosome="3"
/clone="K7M2"
/clone_lib="Mitsui TAC"
complement(1077..2201)
/note="gb|AA03494.1
gene_id:K7M2.2
similar to unknown protein"
/codon_start=1
/evidence=not experimental
/protein_id="BAB02925.1"
/db_xref="GI:11994687"
/translation="MADADELMTKNGVKVTSQSSVQTLRSSVLRLLNPEITRELRRLK
DNNLRKIDHDEKPEENPILNPKMIPRIYIVKPKDDVRKMKDLKQKLTLL
KAIBERLTRDELRYVTCIKDFKSHRIDYQFVKSLQRLIDKYNLYQRFVDIATGHG
EDAGTFVKVDNPDQTESDKIKLEGGKGMPLPKESDOPERLKKRKRTERATPNY
KLIPKSEQTPVSTVLNWTWVNSYDIAQKNLTDIEKMYNHEDQMFELMLVGLFT
SAAKNAEIVNGERDLKLGKFKYCAENIDGRMDLEIVKENHORVLPAILNRLNQKL
REVTLVRELRKPVLKQTIKLTSTRQGSTAQESQVKTRTI"
complement(join(3781..4009,4102..4277))
/note="gene_id:K7M2.3
unknown protein"
/codon_start=1
/evidence=not experimental
/protein_id="BAB02926.1"
/db_xref="GI:11994688"
/translation="MRMVQLLFKLIQVGFVLSNEIGNKNLLLIHHAGSLLS
MFGNDCLIAAPDKKGLTLIASEFSDYFENLVKDRVGNVYVQRLINGFKTKTGIDLPHS
LTSVLVTRSHLKHRYQVIEAFDRSTRLA"
complement(6378..6587)
/note="gene_id:K7M2.4
unknown protein"
/codon_start=1
/evidence=not experimental
/protein_id="BAB02927.1"
/db_xref="GI:11994689"
/translation="MRPMVYRDFLGMEYFGYSMPMQMEIDEDDFGPMDMQFVEYGGIS
PFQMKPDSDFPNKFEDDFDSDIN"
complement(8640..10130)
/note="gene_id:K7M2.5"
/codon_start=1
/evidence=not experimental
/product="ammonium transporter"
/protein_id="BAB02928.1"
/db_xref="GI:11994690"
/translation="MSGATCSAADLSALLGNPATAADYICQGLGSVNNKFTDAAYA
IDNTYLLFSAYLVFMQLGPMILCAGSVRAKNTNMLTNVLDAAAGGLFYILFGYAF
AFGSSDGTGRHFNGLQNPFTLTSYFFLYQWFAIAAAGTSIAERTKFAVYL
IYSSFTGTGVYVSVHFWSPDGMWSPFRSEDLFGTGAIDFAGSGVHVMVGGLW
GALIEGRIGRFPDGGHATALRGHSASLVGLTFLWFGWYGNPQSFTKLIPYNSG
SVYQSGSIGRTAVTTLSCTAAATLTFGKRLLSGHWNVTDVNCGLGSPFAITAGC
SNYDWAIVCGFVASVLIGCNKAELIKYDPLEAAQHLGGCGAGGLIFVGLFAKE
KYINEYVAGSPGRHYGLFMWGGKGLQAQLVQIIVVGVWSATMGTLTFLIKKLNILR
ISQHEMRGMDLAGHGCFAYIYHNDNDDDSIGVPGSPVPAPNPAPV"
complement(13098..14594)
/note="gb|AAD54638.1
gene_id:K7M2.6"
/codon_start=1
/evidence=not experimental
/product="ammonium transporter"
/protein_id="BAB02929.1"
/db_xref="GI:11994691"
/translation="MSGATCSAADLTLGPNATAADYICQGLGSVNNKFTDAAYA
IDNTYLLFSAYLVFMQLGPMILCAGSVRAKNTNMLTNVLDAAAGGLFYILFGYAF
AFGSSDGTGRHFNGLQNPFTLTSYFFLYQWFAIAAAGTSIAERTKFAVYL
IYSSFTGTGVYVSVHFWSPDGMWSPFRSADRLFTGTGAIDFAGSGVHVMVGGLW
WGALEGRPRGRFEKGRALALRGHSASLVGLTFLWFGWYGNPQSFTKLIPYNSG
GSNYQWSIGRTAVNTLSGCTAALTTLFGKRLLSGHWNVTDVNCGLGSPFAITAG

CDS

CDS

CDS

CDS

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 12, 2004, 11:40:28 ; Search time 742 Seconds
(without alignments)
11130.050 Million cell updates/sec

Title: US-10-086-464-1
Perfect score: 1944
Sequence: 1 atgtctcggcgccgtctcc.....atagtggacctctctttaa 1944

Scoring table: OLIGO_NUC
Gapop 60.0, Gapext 60.0

Searched: 3373863 seqs, 2124099041 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N Geneseq_29Jan04:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002s:*
- 7: Geneseqn2003as:*
- 8: Geneseqn2003bs:*
- 9: Geneseqn2003cs:*
- 10: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1944	100.0	1944	4	AAF77094
2	1944	100.0	1944	9	ACF36548
3	1944	100.0	2189	9	ACF36549
4	66	3.4	1959	9	ACF36557
5	47	2.4	2324	3	AAC36968
6	26	1.3	1515	9	ACF36555
7	26	1.3	110000	3	AAF22305_06
8	26	1.3	110000	3	AAF22305_07
9	24	1.2	1389	7	ADA70853
10	24	1.2	1515	8	ADA27371
11	24	1.2	1515	8	ADA92549
12	24	1.2	1515	8	ADA14298
13	24	1.2	1515	8	ADA58571
14	23	1.2	897	7	AA47597
15	23	1.2	897	7	AA47597
16	23	1.2	897	7	AA47598
17	23	1.2	1005	7	ACF70783
18	23	1.2	1281	7	AA47599
19	23	1.2	1572	5	AA584463
20	23	1.2	3252	9	ADB61530
21	23	1.2	4173	9	ADB61585
22	23	1.2	4517	6	ABV99385
23	23	1.2	4586	7	AA48130

24	23	1.2	4620	7	AA47594	Aad47594 Human hep	
25	23	1.2	4620	9	ADC78858	ADC78858 Human PRO	
26	23	1.2	4626	2	AAV20444	AAV20444 Human c-m	
27	23	1.2	110000	7	ACF67367_38	Continuation (39 o	
c	28	1.2	110000	7	ACF65388_09	Continuation (10 o	
29	21	1.1	1119	6	AAQ30797	Aad30797 Hansenula	
30	21	1.1	1269	2	AAQ47060	Aad47060 ZmPK1 hom	
31	21	1.1	1949	4	AAF44708	Abf44708 Novel pro	
32	21	1.1	2048	6	ABN59741	Abn59741 Novel hum	
33	21	1.1	2880	4	AAF77095	Abf77095 Arabidops	
34	21	1.1	2880	9	ACF36550	ACf36550 Arabidops	
35	21	1.1	2891	5	AA07043	Acd07043 Human DNA	
c	36	21	1.1	50959	3	AAF22304	AAF22304 Arabidops
c	37	21	1.1	72592	3	AAF22300	AAF22300 BAC conta
38	21	1.1	222880	9	ADC87622	ADC87622 Human GPC	
39	21	1.1	349938	9	ADC87621	ADC87621 Human GPC	
c	40	20	1.0	213	6	ABL86637	Ab186637 Human ova
41	20	1.0	432	7	ABX62007	Abx62007 Arabidops	
42	20	1.0	609	4	AAF97928	Abf97928 Human sec	
c	43	20	1.0	657	9	ADB95095	Abd95095 A. thalia
c	44	20	1.0	772	3	AAC49631	Aac49631 Arabidops
45	20	1.0	1002	6	AA44411	Acd44411 Human hun	

ALIGNMENTS

RESULT 1

ID	AAF77094	standard; DNA; 1944 BP.
AC	AAF77094;	
XX		
DT	17-MAY-2001	(first entry)
XX		
DE	Brassica napus PERK1 DNA.	
XX		
KW	Proline-rich extensin-like receptor kinase; PERK; resistance; plant; ds.	
XX		
OS	Brassica napus.	
XX		
FN	WO200114563-A1.	
XX		
PD	01-MAR-2001.	
XX		
PF	18-AUG-2000; 2000WO-CA000966.	
XX		
PR	19-AUG-1999; 99US-0149466P.	
PR	13-OCT-1999; 99US-0159122P.	
XX		
PA	(GORI/) GORING D.	
PA	(SILV/) SILVA N.	
XX		
PI	Goring D, Silva N;	
XX		
DR	WPI; 2001-244305/25.	
XX		
PT	New proline-rich, extensin-like receptor kinase nucleic acids and	
PT	polypeptides useful for increasing plant wounding or pathogen resistance,	
PT	or for producing transgenic plants with increased wounding or pathogen	
PT	resistance.	
XX		
PS	Claim 6; Fig 1; 91pp; English.	
XX		
CC	The present invention relates to proline-rich extensin-like receptor	
CC	kinase (PERK). The PERK nucleic acids and polypeptides are useful for	
CC	increasing the resistance of plants to wounding and pathogens. These are	
CC	also useful for producing transgenic plants with increased wounding and	
CC	pathogen resistance compared with a wild type plant, as well as in assays	
CC	for identifying and developing compounds to inhibit and/or enhance	
XX	polypeptide function directly	
SQ	Sequence 1944 BP; 464 A; 550 C; 459 G; 471 T; 0 U; 0 Other;	

DE	B. napus PERK1 receptor kinase encoding cDNA.	
XX	PERK1; transgenic; plant; proline-rich extensin-like receptor kinase;	
KW	wound; pathogen resistance; plant growth; seed production; gene; ss.	
XX	Brassica napus.	
OS		
XX	Key Location/Qualifiers	
XX	CDS 1..1944	
FT	/*tag= a	
FT	/transl_except= (pos: 1332..1334, aa: Phe)	
FT	/product= "PERK1"	
FT	/note= "Proline-rich Extensin-like Receptor Kinase"	
XX	MO2003072763-A1.	
PN		
XX	04-SEP-2003.	
PD		
XX	28-FEB-2003; 2003WO-CA000274.	
XX	28-FEB-2002; 2002CA-02373903.	
PR	28-FEB-2002; 2002US-00086464.	
PR		
XX	(GORI/) GORING D.	
PA	(SILV/) SILVA N.	
PA	(HAFF/) HAFFANI Y Z.	
XX		
XX	Goring D, Silva N, Haffani YZ;	
PI		
XX	WPI; 2003-712727/67.	
DR	P-PSDB; ABR82937.	
XX		
DR	Producing a transgenic plant having an increased plant resistance, plant	
PT	growth or seed production comprises transforming a plant with a nucleic	
PT	acid molecule having a proline-rich Extensin-like Receptor Kinase	
PT	activity.	
XX		
XX	Claim 3; Fig 1A; 123pp; English.	
XX		
CC	The invention relates to producing a transgenic plant having increased	
CC	plant height, number of branches, number of seed pods and/or seed	
CC	production compared to a non-transgenic plant, and/or quicker flowering	
CC	or later senescence compared to a non-transgenic plant. The method	
CC	involves transforming a plant with a vector including a proline-rich	
CC	Extensin-like Receptor Kinase (PERK) nucleic acid molecule or a nucleic	
CC	acid molecule having PERK activity. The method, as well as the PERK	
CC	nucleic acid molecule and polypeptide, are useful in increasing plant	
CC	resistance to wounding and pathogens and in increasing plant growth and	
CC	seed production. The nucleic acid molecule and polypeptide may also be	
CC	used in producing transgenic plants or transgenic host cells. The present	
CC	sequence represents a cDNA encoding a B. napus PERK1 receptor kinase	
CC	polypeptide	
XX		
XX	Sequence 1944 BP; 464 A; 550 C; 459 G; 471 T; 0 U; 0 Other;	
QY	Query Match 100.0%; Score 1944; DB 9; Length 1944;	
DB	Best Local Similarity 100.0%; Pred. No. 0;	
DB	Matches 1944; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 ATGTCTCTCGGCGCGTCTCGGGGACTGGTTCGCTCCATCTCCACCATCAAACTCCACA 60	
DB	1 ATGTCTCTCGGCGCGTCTCGGGGACTGGTTCGCTCCATCTCCACCATCAAACTCCACA 60	
QY	61 ACCACCACTCTCTCAGCTTCGGCTCTCTCTCCGACACACCTTCTCTCTCGCGCG 120	
DB	61 ACCACCACTCTCTCAGCTTCGGCTCTCTCTCTCCGACACACCTTCTCTCTCGCGCG 120	
QY	121 CCATCCACTATCCGACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180	
DB	121 CCATCCACTATCCGACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180	
QY	181 CCATCTCCACCACTCCATCTACCGCGGATCTCCACCTCTCTCTCTCTCTCTCTCTCCA 240	
DB	181 CCATCTCCACCACTCCATCTACCGCGGATCTCCACCTCTCTCTCTCTCTCTCTCTCCA 240	

Db	181	CCATCTCCACCACTCCATCTACCGCGGATCTCCACCTCTCTCTCTCTCTCTCTCTCCA	240
QY	241	CCCGCTCCAACTACGCGCGGATCTCCACCGCGACCTGTGTACTCTCTCTCTCTCTCTCT	300
Db	241	CCCGCTCCAACTACGCGCGGATCTCCACCGCGACCTGTGTACTCTCTCTCTCTCTCTCT	300
QY	301	CCACCTTTCAGTCCAGGACCACTCTCAATCTCTTACGCGAAGAGGATCTCTCTCGACCT	360
Db	301	CCACCTTTCAGTCCAGGACCACTCTCAATCTCTTACGCGAAGAGGATCTCTCTCGACCT	360
QY	361	CCATCT	420
Db	361	CCATCT	420
QY	421	GGAAATCGCCATCGAGGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	480
Db	421	GGAAATCGCCATCGAGGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	480
QY	481	AAGAAGAAACGACGAGAGACGAAAGATGTCTTACTATGTTCCTCGCCACCTCTCTCTCT	540
Db	481	AAGAAGAAACGACGAGAGACGAAAGATGTCTTACTATGTTCCTCGCCACCTCTCTCTCT	540
QY	541	GGTCCCAAAGCCGAGGACCTTACGGTGGACAGCAACAATGGCGGCAACAAACGCA	600
Db	541	GGTCCCAAAGCCGAGGACCTTACGGTGGACAGCAACAATGGCGGCAACAAACGCA	600
QY	601	ACACCACTCTCAGATCATGTCTGAGTCACTACCACTACCACTACCACTACCACTACCA	660
Db	601	ACACCACTCTCAGATCATGTCTGAGTCACTACCACTACCACTACCACTACCACTACCA	660
QY	661	CCACGCACT	720
Db	661	CCACGCACT	720
QY	721	TACTCGGACCGTCCAGTCT	780
Db	721	TACTCGGACCGTCCAGTCT	780
QY	781	AGACATTTTACATACGAGGAGTCTAGAGCCACCAATGGTTCCTCCGAGGCGAACTTG	840
Db	781	AGACATTTTACATACGAGGAGTCTAGAGCCACCAATGGTTCCTCCGAGGCGAACTTG	840
QY	841	TTAGGCAACAGCGCGGTTCGGTTACGTGCACAAAGTGTGTGTCTAGTGGGAAAGAGTT	900
Db	841	TTAGGCAACAGCGCGGTTCGGTTACGTGCACAAAGTGTGTGTCTAGTGGGAAAGAGTT	900
QY	901	GCTGTGAAGCAGTGTGAAGTTGGGAGTGTGTCAGGAGAGAGGAGTTTCAGGAGAGGTT	960
Db	901	GCTGTGAAGCAGTGTGAAGTTGGGAGTGTGTCAGGAGAGAGGAGTTTCAGGAGAGGTT	960
QY	961	GAGATCATCAGCAGAGTTCACCACTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1020
Db	961	GAGATCATCAGCAGAGTTCACCACTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1020
QY	1021	GGTGCACAAAGATTTGCTTGTATGATTTGTTTCTTAAACCAATCTCCAGCTTCACCTC	1080
Db	1021	GGTGCACAAAGATTTGCTTGTATGATTTGTTTCTTAAACCAATCTCCAGCTTCACCTC	1080
QY	1081	CATGGCGAGGACGGCTTCAATGGAATGGAGCACCAGATTTGAAGATTTGCTCTTGGATCT	1140
Db	1081	CATGGCGAGGACGGCTTCAATGGAATGGAGCACCAGATTTGAAGATTTGCTCTTGGATCT	1140
QY	1141	GCTAAAGGACTTTCTTATCTTCAATGGAATTTCAAGTTTGAAGCTTGAAGTTTGGTCTCT	1200
Db	1141	GCTAAAGGACTTTCTTATCTTCAATGGAATTTCAAGTTTGAAGCTTGAAGTTTGGTCTCT	1200
QY	1201	AAGGCTTCAAAACATATTTAGATTTTCAAGTTTGAAGCTTGAAGTTTGGTCTCT	1260
Db	1201	AAGGCTTCAAAACATATTTAGATTTTCAAGTTTGAAGCTTGAAGTTTGGTCTCT	1260
QY	1261	GCTAAGATTTCTCTGATCAAAACAGCATATTAACACGCTGTGATGGGAACTTTGGG	1320
Db	1261	GCTAAGATTTCTCTGATCAAAACAGCATATTAACACGCTGTGATGGGAACTTTGGG	1320

```
Qy 1321 TACTTGGCTCCGGAATACGCTGCAAGCGGAAGCTCACGGAGAAAGTCTGACGTTTTCTCA 1380
Db 1321 TACTTGGCTCCGGAATACGCTGCAAGCGGAAGCTCACGGAGAAAGTCTGACGTTTTCTCA 1380
Qy 1381 TTTGGCGTTGGCTTTTGGAGCTCATTACTGGAGCGTGCACCGCTTGATGCCAACAATGTC 1440
Db 1381 TTTGGCGTTGGCTTTTGGAGCTCATTACTGGAGCGTGCACCGCTTGATGCCAACAATGTC 1440
Qy 1441 TATGTAGATGACAGCTTAGTTGACTGGGCACGACCATTTCTTAACCGAGCATCTGAGCAA 1500
Db 1441 TATGTAGATGACAGCTTAGTTGACTGGGCACGACCATTTCTTAACCGAGCATCTGAGCAA 1500
Qy 1501 GGAGACTTTGAGGGTTAGCTGATGCAAAAGATGAATAATGGGTATGACAGAGAGGATG 1560
Db 1501 GGAGACTTTGAGGGTTAGCTGATGCAAAAGATGAATAATGGGTATGACAGAGAGGATG 1560
Qy 1561 GCTGCGATGTTGCTTGTGCTGCGGCTTGTTCGCCCATTCAGCTCGCGCGAGACCTCGC 1620
Db 1561 GCTGCGATGTTGCTTGTGCTGCGGCTTGTTCGCCCATTCAGCTCGCGCGAGACCTCGC 1620
Qy 1621 ATGAGCCAGATTGTCGTCGCTTAGAAGGAATGTATCACTGTCTGATCTTTAAACGAAGGG 1680
Db 1621 ATGAGCCAGATTGTCGTCGCTTAGAAGGAATGTATCACTGTCTGATCTTTAAACGAAGGG 1680
Qy 1681 ATGAGACAGCTCAAGCAATGTATACAGCTCATACGGAGGAACCGATTTATGACTCG 1740
Db 1681 ATGAGACAGCTCAAGCAATGTATACAGCTCATACGGAGGAACCGATTTATGACTCG 1740
Qy 1741 AGCCAGTACATGAAGACATGAAGATTTTAGGAAATGGCACTTGGAACTCAAGAGTAC 1800
Db 1741 AGCCAGTACATGAAGACATGAAGATTTTAGGAAATGGCACTTGGAACTCAAGAGTAC 1800
Qy 1801 AACCGCCAGGTGAGTACAGTAATCCGACCAGTGTATGACGTATGGACTGTACCCGCTCTGTTCA 1860
Db 1801 AACCGCCAGGTGAGTACAGTAATCCGACCAGTGTATGACGTATGGACTGTACCCGCTCTGTTCA 1860
Qy 1861 AGCAGCAGGCGCCAAACACACGCGAAATGAGATGGGGAAGATTAAGAGAACCGGTGAG 1920
Db 1861 AGCAGCAGGCGCCAAACACACGCGAAATGAGATGGGGAAGATTAAGAGAACCGGTGAG 1920
Qy 1921 GGTATAGTGACCTTCTCTTTAA 1944
Db 1921 GGTATAGTGACCTTCTCTTTAA 1944
```

```
RESULT 3
ACF36549
ID ACF36549 standard; cdna; 2189 BP.
```

```
XX ACF36549;
```

```
DT 18-DEC-2003 (first entry)
```

```
DE B. napus PERK1 receptor kinase encoding cdna.
```

```
XX PERK1; transgenic; plant; proline-rich extensin-like receptor kinase;
XX wound; pathogen resistance; plant growth; seed production; gene; ss.
```

```
OS Brassica napus.
```

```
XX Key Location/Qualifiers
FH 1. .2189
CDS /*tag= c
```

```
FT /note= "ABR82938; this protein contains the amino acids
FT corresponding to 5' and 3'UTR regions, though only the
FT relevant aa residues (ABR82937) is used in the invention"
```

```
FT 5'UTR
```

```
FT 1. .96
```

```
FT /*tag= a
```

```
FT 97. .2040
```

```
FT /*tag= b
```

```
FT /transl_except= (pos: 1332. .1334, aa: Phe)
```

```
FT /product= "PERK1"
```

```
FT /note= "ABR82937; proline-rich Extensin-like Receptor
FT Kinase"
FT 3'UTR 2041. .2189
FT /*tag= d
```

```
XX W02003072763-A1.
```

```
XX 04-SEP-2003.
```

```
XX 28-FEB-2003; 2003WO-CA000274.
```

```
XX 28-FEB-2002; 2002CA-02373903.
```

```
XX 28-FEB-2002; 2002US-00086464.
```

```
XX (GORI/) GORING D.
```

```
XX (SILV/) SILVA N.
```

```
XX (HAFF/) HAFFANI Y Z.
```

```
XX Goring D, Silva N, Haffani YZ;
```

```
XX WPI; 2003-712727/67.
```

```
XX P-PSDB; ABR82937, ABR82938.
```

```
XX Producing a transgenic plant having an increased plant resistance, plant
XX growth or seed production comprises transforming a plant with a nucleic
XX acid molecule having a Proline-rich Extensin-like Receptor Kinase
XX activity.
```

```
XX Disclosure; Fig 1D; 123pp; English.
```

```
XX The invention relates to producing a transgenic plant having increased
XX plant height, number of branches, number of seed pods and/or seed
XX production compared to a non-transgenic plant, and/or quicker flowering
XX or later senescence compared to a non-transgenic plant. The method
XX involves transforming a plant with a vector including a Proline-rich
XX Extensin-like Receptor Kinase (PERK) nucleic acid molecule or a nucleic
XX acid molecule having PERK activity. The method, as well as the PERK
XX nucleic acid molecule and polypeptide, are useful in increasing plant
XX resistance to wounding and pathogens and in increasing plant growth and
XX seed production. The nucleic acid molecule and polypeptide may also be
XX used in producing transgenic plants or transgenic host cells. The present
XX sequence represents a cdna encoding a B. napus PERK1 receptor kinase
XX polypeptide
```

```
XX Sequence 2189 BP; 544 A; 592 C; 494 G; 559 T; 0 U; 0 Other;
```

```
Query Match 100.0%; Score 1944; DB 9; Length 2189;
```

```
Best Local Similarity 100.0%; Pred. No. 0;
```

```
Matches 1944; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 ATGTCTCTGGCGCGTCTCCGGGACTGGTTCGCTCCATCTCCACCATCAAACTCCACA 60
Db 97 ATGTCTCTGGCGCGTCTCCGGGACTGGTTCGCTCCATCTCCACCATCAAACTCCACA 156
```

```
Qy 61 ACACGACTCTCTCTCCAGCTTCGGCTCTCTCCGACGACACCTTCTCTCTCCGCGG 120
Db 157 ACCACCACTCTCTCTCCAGCTTCGGCTCTCTCCGACGACACCTTCTCTCTCCGCGG 216
```

```
Qy 121 CCATCCACTATTCGACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180
Db 217 CCATCCACTATTCGACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 276
```

```
Qy 181 CCATCTCCACCACTCCATCTACGCGGGATCTCCACCTCTCTCTCTCTCTCTCTCTCCA 240
Db 277 CCATCTCCACCACTCCATCTACGCGGGATCTCCACCTCTCTCTCTCTCTCTCTCTCCA 336
```

```
Qy 241 CCGCTCTCAACTAGCCCGGATCTCCACCGGACCTCTTACTCTCTCTCTCTCTCTCTCT 300
Db 337 CCGCTCTCAACTAGCCCGGATCTCCACCGGACCTCTTACTCTCTCTCTCTCTCTCTCT 396
```

```
Qy 301 CCACCTTCAGTCCCGAGGACCAACCGTCCCAATCTCTTACGCGAAGGAGGATCTCTCTCGACCT 360
Db 397 CCACCTTCAGTCCCGAGGACCAACCGTCCCAATCTCTTACGCGAAGGAGGATCTCTCTCGACCT 456
```

QY 361 CCATCTTCTCCCTCGCGCGTCTCTCTTCTCCGACGGTTTATCAACAGGAGTGGTGGT 420
DB 457 CCATCTTCTCCCTCGCGCGTCTCTCTTCTCCGACGGTTTATCAACAGGAGTGGTGGT 516
QY 421 GGAATCCCATCGGAGAGTGGTCTCTCTTGTGATGACTCTGATTTGTCTCTCTGT 480
DB 517 GGAATCCCATCGGAGAGTGGTCTCTCTTGTGATGACTCTGATTTGTCTCTCTGT 576
QY 481 AAGNAGAAACGACGGAGAGAGAGAGATGCTTACTATGTTCTCTCGCCACCTCTCTCT 540
DB 577 AAGNAGAAACGACGGAGAGAGAGAGATGCTTACTATGTTCTCTCGCCACCTCTCTCT 636
QY 541 GGTCCCAAGCGGAGAGCTTACGGTGGACAGCAACAAATGGCGGCAACAAACGCA 600
DB 637 GGTCCCAAGCGGAGAGCTTACGGTGGACAGCAACAAATGGCGGCAACAAACGCA 696
QY 601 ACACACCGTCAGATCATGTCGTGACGTCACTACCAACCACTTAAGGTCCATCTCCA 660
DB 697 ACACACCGTCAGATCATGTCGTGACGTCACTACCAACCACTTAAGGTCCATCTCCA 756
QY 661 CCAGGCAACCTCTCTCACTCCACCGCTCTCCAGCGCTTGTGTTAGGCTTCTCCAAA 720
DB 757 CCAGGCAACCTCTCTCACTCCACCGCTCTCCAGCGCTTGTGTTAGGCTTCTCCAAA 816
QY 721 TACTCGGACCGTCCAGTCTCTCTCCACCGCTCTCCAGCGCTTGTGTTAGGCTTCTCCAAA 780
DB 817 TACTCGGACCGTCCAGTCTCTCTCCACCGCTCTCCAGCGCTTGTGTTAGGCTTCTCCAAA 876
QY 781 AGCATTTTCAATACGAGGAGTCTAGTCTAGAGCCCAATGGTTTCTCCAGGCGCACTTG 840
DB 877 AGCATTTTCAATACGAGGAGTCTAGTCTAGAGCCCAATGGTTTCTCCAGGCGCACTTG 936
QY 841 TTAGGACAAAGCGGTTTCGGTTAGTCTGCAAAAGGTGTGTCCTAGTGGGAAAGAGTT 900
DB 937 TTAGGACAAAGCGGTTTCGGTTAGTCTGCAAAAGGTGTGTCCTAGTGGGAAAGAGTT 996
QY 901 GCTGTGAAGCAGTTGAAGTTGGAGTGGTCAGGAGAGAGGAGTTTTCAGGACAGAGTT 960
DB 997 GCTGTGAAGCAGTTGAAGTTGGAGTGGTCAGGAGAGAGGAGTTTTCAGGACAGAGTT 1056
QY 961 GAGATCATACGAGAGTTTCAACACAGGATCTGGTGTCTCTTGTGTTATTTGATCGCC 1020
DB 1057 GAGATCATACGAGAGTTTCAACACAGGATCTGGTGTCTCTTGTGTTATTTGATCGCC 1116
QY 1021 GGTGCGAAAGATTGCTTGTCTATGATTTGTTCTTAACAAATCTCGAGCTTCACCTC 1080
DB 1117 GGTGCGAAAGATTGCTTGTCTATGATTTGTTCTTAACAAATCTCGAGCTTCACCTC 1176
QY 1081 CATGGCGAGGAGCGGCTCAATGGAATGGAGCACCAGATTGAAGATTGCTCTTGGATCT 1140
DB 1177 CATGGCGAGGAGCGGCTCAATGGAATGGAGCACCAGATTGAAGATTGCTCTTGGATCT 1236
QY 1141 GCTAAAGGATCTTCTTATCTTCAATGAAGATTGCAATCTTAAATCAATTCACCGTGATATC 1200
DB 1237 GCTAAAGGATCTTCTTATCTTCAATGAAGATTGCAATCTTAAATCAATTCACCGTGATATC 1296
QY 1201 AAGCTTTCAACATATTTGATGATTTCAAGTTTGAAGCTAAGGTTGCTGATTTTGGTCTT 1260
DB 1297 AAGCTTTCAACATATTTGATGATTTCAAGTTTGAAGCTAAGGTTGCTGATTTTGGTCTT 1356
QY 1261 GCTAAGATTGCTTCTGATCAACACCGCATGTATCAACAGTGTGATGGGAACCTTTGGG 1320
DB 1357 GCTAAGATTGCTTCTGATCAACACCGCATGTATCAACAGTGTGATGGGAACCTTTGGG 1416
QY 1321 TACTTGGCTCGGAATACGCTCAAGCGGAAAGCTCACGGAGAGTCTGACGTTTCTCA 1380
DB 1417 TACTTGGCTCGGAATACGCTCAAGCGGAAAGCTCACGGAGAGTCTGACGTTTCTCA 1476
QY 1381 TTTGGCTTGTCTTTTGGAGCTCATTTAGGAGTGCACCCGTTGATGCCAATATGTC 1440
DB 1477 TTTGGCTTGTCTTTTGGAGCTCATTTAGGAGTGCACCCGTTGATGCCAATATGTC 1536

QY 1441 TATGTAGATGACAGCTTAGTGTGCTGGCCACGACCACTTGTCTTAACCGAGCATCTGACAA 1500
DB 1537 TATGTAGATGACAGCTTAGTGTGCTGGCCACGACCACTTGTCTTAACCGAGCATCTGACAA 1596
QY 1501 GAGACTTTGAGGGTTAGCTGATGCAAAAGATGAATTAATGGGTATGACAGAGAGGATG 1560
DB 1597 GAGACTTTGAGGGTTAGCTGATGCAAAAGATGAATTAATGGGTATGACAGAGAGGATG 1656
QY 1561 GCTCGCATGGTTCCTGCTGCTGGCTTGTGTCGCCATTCAGCTCCGCCGAGACCTCGC 1620
DB 1657 GCTCGCATGGTTCCTGCTGCTGGCTTGTGTCGCCATTCAGCTCCGCCGAGACCTCGC 1716
QY 1621 ATGAGCCAGATTGCTGGCTGTAGAGGAAATGTATCACTGTCAAGATCTTAACGAAAGG 1680
DB 1717 ATGAGCCAGATTGCTGGCTGTAGAGGAAATGTATCACTGTCAAGATCTTAACGAAAGG 1776
QY 1681 ATGAGCCAGATTGCTGGCTGTAGAGGAAATGTATCACTGTCAAGATCTTAACGAAAGG 1740
DB 1777 ATGAGCCAGATTGCTGGCTGTAGAGGAAATGTATCACTGTCAAGATCTTAACGAAAGG 1836
QY 1741 AGCCAGTACAAATGAAGACATGAAGAGTTTGTAGGAAATGGCACTTGAAGCTCAAGAGTAC 1800
DB 1837 AGCCAGTACAAATGAAGACATGAAGAGTTTGTAGGAAATGGCACTTGAAGCTCAAGAGTAC 1896
QY 1801 AACGCCACGGGTGAGTACAGTAATCCGACCACTGACTATGAGACTGTACCGGTCTGGTTCA 1860
DB 1897 AACGCCACGGGTGAGTACAGTAATCCGACCACTGACTATGAGACTGTACCGGTCTGGTTCA 1956
QY 1861 AGCAGCGAGGGCCAAACCAACACCGGAAATGGAGATGGGAAAGATTAAGAGAACCGGTGAG 1920
DB 1957 AGCAGCGAGGGCCAAACCAACACCGGAAATGGAGATGGGAAAGATTAAGAGAACCGGTGAG 2016
QY 1921 GGTATAGTGGACCTTCTCTTTAA 1944
DB 2017 GGTATAGTGGACCTTCTCTTTAA 2040

RESULT 4

ACF36557
ID ACF36557 standard; DNA; 1959 BP.
XX ACF36557;
AC AC
XX XX
DT 18-DEC-2003 (first entry)
XX XX
DE A. thaliana PERK1 protein encoding genomic DNA.
XX PERK1; transgenic; plant; proline-rich extensin-like receptor kinase;
KW wound; pathogen resistance; plant growth; seed production; gene; ds.
XX Arabidopsis thaliana.
OS Arabidopsis thaliana.
XX WO2003072763-A1.
PN 04-SEP-2003.
XX 28-FEB-2003; 2003WO-CA000274.
PF 28-FEB-2002; 2002CA-02373903.
XX 28-FEB-2002; 2002US-00086464.
PR (GRI/) GORING D.
PA (SILV/) SILVA N.
XX (HAFF/) HAFFANI Y Z.
PI Goring D, Silva N, Haffani YZ;
XX WPI; 2003-712727/67.
XX Producing a transgenic plant having an increased plant resistance, plant
PT growth or seed production comprises transforming a plant with a nucleic
PT acid molecule having a Proline-rich Extensin-like Receptor Kinase
PT activity.

PR 22-JUL-1999; 99US-01451192P.
 PR 23-JUL-1999; 99US-01451145P.
 PR 23-JUL-1999; 99US-0145218P.
 PR 23-JUL-1999; 99US-0145218P.
 PR 26-JUL-1999; 99US-0145224P.
 PR 26-JUL-1999; 99US-0145226P.
 PR 27-JUL-1999; 99US-0145913P.
 PR 27-JUL-1999; 99US-0145918P.
 PR 27-JUL-1999; 99US-0145918P.
 PR 28-JUL-1999; 99US-0145951P.
 PR 02-AUG-1999; 99US-0146386P.
 PR 02-AUG-1999; 99US-0146388P.
 PR 02-AUG-1999; 99US-0146389P.
 PR 03-AUG-1999; 99US-0147038P.
 PR 04-AUG-1999; 99US-0147204P.
 PR 04-AUG-1999; 99US-0147302P.
 PR 05-AUG-1999; 99US-0147192P.
 PR 06-AUG-1999; 99US-0147260P.
 PR 06-AUG-1999; 99US-0147303P.
 PR 06-AUG-1999; 99US-0147416P.
 PR 09-AUG-1999; 99US-0147493P.
 PR 09-AUG-1999; 99US-0147935P.
 PR 10-AUG-1999; 99US-0148117P.
 PR 11-AUG-1999; 99US-0148319P.
 PR 12-AUG-1999; 99US-0148341P.
 PR 13-AUG-1999; 99US-0148565P.
 PR 13-AUG-1999; 99US-0148694P.
 PR 16-AUG-1999; 99US-0149368P.
 PR 17-AUG-1999; 99US-0149175P.
 PR 18-AUG-1999; 99US-0149426P.
 PR 20-AUG-1999; 99US-0149722P.
 PR 20-AUG-1999; 99US-0149723P.
 PR 20-AUG-1999; 99US-0149929P.
 PR 23-AUG-1999; 99US-0149902P.
 PR 23-AUG-1999; 99US-0149930P.
 PR 23-AUG-1999; 99US-0150566P.
 PR 26-AUG-1999; 99US-0150884P.
 PR 27-AUG-1999; 99US-0151065P.
 PR 27-AUG-1999; 99US-0151066P.
 PR 27-AUG-1999; 99US-0151080P.
 PR 30-AUG-1999; 99US-0151303P.
 PR 31-AUG-1999; 99US-0151438P.
 PR 01-SEP-1999; 99US-0151930P.
 PR 07-SEP-1999; 99US-0152363P.
 PR 10-SEP-1999; 99US-0153070P.
 PR 13-SEP-1999; 99US-0153758P.
 PR 15-SEP-1999; 99US-0154018P.
 PR 16-SEP-1999; 99US-0154039P.
 PR 20-SEP-1999; 99US-0154779P.
 PR 22-SEP-1999; 99US-0155139P.
 PR 23-SEP-1999; 99US-0155486P.
 PR 24-SEP-1999; 99US-0155659P.
 PR 28-SEP-1999; 99US-0156458P.
 PR 29-SEP-1999; 99US-0156596P.
 PR 04-OCT-1999; 99US-0157117P.
 PR 05-OCT-1999; 99US-0157753P.
 PR 06-OCT-1999; 99US-0157865P.
 PR 07-OCT-1999; 99US-0158029P.
 PR 08-OCT-1999; 99US-0158232P.
 PR 12-OCT-1999; 99US-0158369P.
 PR 13-OCT-1999; 99US-0159293P.
 PR 13-OCT-1999; 99US-0159294P.
 PR 13-OCT-1999; 99US-0159295P.
 PR 14-OCT-1999; 99US-0159329P.
 PR 14-OCT-1999; 99US-0159330P.
 PR 14-OCT-1999; 99US-0159331P.
 PR 14-OCT-1999; 99US-0159637P.
 PR 14-OCT-1999; 99US-0159638P.
 PR 18-OCT-1999; 99US-0159584P.
 PR 21-OCT-1999; 99US-0160741P.
 PR 21-OCT-1999; 99US-0160767P.
 PR 21-OCT-1999; 99US-0160768P.
 PR 21-OCT-1999; 99US-0160770P.
 PR 21-OCT-1999; 99US-0160814P.

PR 21-OCT-1999; 99US-0160815P.
 PR 22-OCT-1999; 99US-0160980P.
 PR 22-OCT-1999; 99US-0160981P.
 PR 22-OCT-1999; 99US-0160989P.
 PR 23-OCT-1999; 99US-0161404P.
 PR 25-OCT-1999; 99US-0161405P.
 PR 25-OCT-1999; 99US-0161406P.
 PR 26-OCT-1999; 99US-0161359P.
 PR 26-OCT-1999; 99US-0161360P.
 PR 26-OCT-1999; 99US-0161361P.
 PR 28-OCT-1999; 99US-0161920P.
 PR 28-OCT-1999; 99US-0161992P.
 PR 28-OCT-1999; 99US-0161993P.
 PR 29-OCT-1999; 99US-0162142P.

Query Match 2.4%; Score 47; DB 3; Length 2324;

Best Local Similarity 100.0%; Pred. No. 1.7e-11;
 Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1585 GCTTGTTCGCCATTGCGCTCGCCGAGACCTCGCATGAGCCAGAT 1631
 |||||
 Db 1700 GCTTGTTCGCCATTGCGCTCGCCGAGACCTCGCATGAGCCAGAT 1746

RESULT 6

ACF36555
 ID ACF36555 standard; DNA; 1515 BP.

XX ACF36555;

XX 18-DEC-2003 (first entry)

DE Z. mays PERK protein encoding genomic DNA.

KW PERK1; transgenic; plant; proline-rich extensin-like receptor kinase;
 KW wound; pathogen resistance; plant growth; seed production; maize; gene;
 ds.

XX Zea mays.

XX WO2003072763-A1.

XX 04-SEP-2003.

XX 28-FEB-2003; 2003WO-CA000274.

XX 28-FEB-2002; 2002CA-02373903.

XX 28-FEB-2002; 2002US-00086464.

XX (GORI/) GORING D.

XX (SILV/) SILVA N.

XX (HAFF/) HAFFANI Y Z.

XX Goring D, Silva N, Haffani YZ;

XX WPI; 2003-712727/67.

PT Producing a transgenic plant having an increased plant resistance, plant
 growth or seed production comprises transforming a plant with a nucleic
 acid molecule having a Proline-rich Extensin-like Receptor Kinase
 activity.

XX Disclosure; Fig 21b; 123pp; English.

CC The invention relates to producing a transgenic plant having increased
 plant height, number of branches, number of seed pods and/or seed
 production compared to a non-transgenic plant, and/or quicker flowering
 or later senescence compared to a non-transgenic plant. The method
 involves transforming a plant with a vector including a proline-rich
 Extensin-like Receptor Kinase (PERK) nucleic acid molecule or a nucleic
 acid molecule having PERK activity. The method, as well as the PERK
 nucleic acid molecule and polypeptide, are useful in increasing plant
 resistance to wounding and pathogens and in increasing plant growth and

CC seed production. The nucleic acid molecule and polypeptide may also be
CC used in producing transgenic plants or transgenic host cells. The present
CC sequence represents a Z. mays PERK protein encoding genomic DNA (TIGR
CC Accession No. AY108241)

XX Sequence 1515 BP; 400 A; 330 C; 381 G; 404 T; 0 U; 0 Other;

Query Match 1.3%; Score 26; DB 9; Length 1515;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1231 TTTGAAGCTAAGTGTGCTGATTGG 1256

Db 601 TTTGAAGCTAAGTGTGCTGATTGG 626

RESULT 7

AAF22305_06/c

Continuation (7 of 11) of AAF22305 from base 600001 (Arabidopsis thaliana chromosome 4 c
WP Sequence split into 11 fragments LOCUS AAF22305 Accession Aaf22305

WP	Fragment Name	Begin	End
WP	AAF22305_00	1	110000
WP	AAF22305_01	100001	210000
WP	AAF22305_02	200001	310000
WP	AAF22305_03	300001	410000
WP	AAF22305_04	400001	510000
WP	AAF22305_05	500001	610000
WP	AAF22305_06	600001	710000
WP	AAF22305_07	700001	810000
WP	AAF22305_08	800001	910000
WP	AAF22305_09	900001	1010000
WP	AAF22305_10	1000001	1082138

Query Match 1.3%; Score 26; DB 3; Length 110000;

Best Local Similarity 100.0%; Pred. No. 0.1;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1036 CTTGCTATGAGTTGTTCTTAACAA 1061

Db 109643 CTTGCTATGAGTTGTTCTTAACAA 109618

RESULT 8

AAF22305_07/c

Continuation (8 of 11) of AAF22305 from base 700001 (Arabidopsis thaliana chromosome 4 c
WP Sequence split into 11 fragments LOCUS AAF22305 Accession Aaf22305

WP	Fragment Name	Begin	End
WP	AAF22305_00	1	110000
WP	AAF22305_01	100001	210000
WP	AAF22305_02	200001	310000
WP	AAF22305_03	300001	410000
WP	AAF22305_04	400001	510000
WP	AAF22305_05	500001	610000
WP	AAF22305_06	600001	710000
WP	AAF22305_07	700001	810000
WP	AAF22305_08	800001	910000
WP	AAF22305_09	900001	1010000
WP	AAF22305_10	1000001	1082138

Query Match 1.3%; Score 26; DB 3; Length 110000;

Best Local Similarity 100.0%; Pred. No. 0.1;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1036 CTTGCTATGAGTTGTTCTTAACAA 1061

Db 9643 CTTGCTATGAGTTGTTCTTAACAA 9618

RESULT 9

ADA70853/c

ID ADA70853 standard; DNA; 1389 BP.

XX

AC ADA70853;

XX 20-NOV-2003 (first entry)

XX Rice gene, SEQ ID 4176.

XX Plant; bacterial infection; fungal infection; viral infection; rice;
XX gene; ds.

XX Oryza sativa.

XX WO2003000898-A1.

XX 03-JAN-2003.

XX 22-JUN-2001; 2001WO-IB001105.

XX 22-JUN-2001; 2001WO-IB001105.

XX (SYGN) SYNGENTA PARTICIPATIONS AG.

XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
XX Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;

XX WPI; 2003-175290/17.

XX Identifying at least one gene involved in plant resistance or response to
XX pathogenic infection for conferring resistance or tolerance to a plant to
XX bacterial, fungal or viral infection by determining or detecting plant
XX gene expression.

XX Claim 6; SEQ ID NO 4176; 899pp; English.

XX The present invention relates to a method (M1) for identifying genes
XX involved in plant resistance or response to pathogenic infection. M1
XX comprises identifying a gene whose expression is significantly altered in
XX the incompatible interaction of plant gene expression relative to
XX expression of the gene in an uninfected plant, in a mutant plant that
XX does not express a gene associated with response to pathogenic infection,
XX or in a corresponding incompatible or compatible interaction. (M1) is
XX useful for conferring resistance to resistance or tolerance to a plant to
XX bacterial, fungal or viral infection. The present sequence was used to
XX illustrate the invention.

XX Sequence 1389 BP; 412 A; 278 C; 406 G; 293 T; 0 U; 0 Other;

Query Match 1.2%; Score 24; DB 7; Length 1389;

Best Local Similarity 100.0%; Pred. No. 1;

Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 670 CCTCTCCAGCTCCACCACCGCCT 693

Db 115 CCTCTCCAGCTCCACCACCGCCT 92

RESULT 10

ADA27371

ID ADA27371 standard; DNA; 1515 BP.

XX ADA27371;

XX 20-NOV-2003 (first entry)

XX HPV-16 L1 codon optimised DNA sequence SEQ ID NO:1.

XX cell line; American Type Culture Collection PTA-4047; ATCC-4047;
XX baculoviruses; viral recombinant protein; virus-like particle; vaccine;
XX diagnostic reagent; human papillomavirus type 16; HPV-16; L1; gene; ds.

XX Synthetic.

XX Human papillomavirus type 16.

XX Key Location/Qualifiers

XX CDS 1. .1515

```
FT      /*tag= a
FT      /product= "HPV-16 L1 protein"
FT      /transl_except= (pos:208. .210,aa:Tyr)
FT      /transl_except= (pos:295. .297,aa:Val)
XX
XX      WO2003068804-A2.
XX
XX      21-AUG-2003.
XX
XX      14-FEB-2003; 2003WO-US004516.
XX
XX      14-FEB-2002; 2002US-0356113P.
XX      14-FEB-2002; 2002US-0356118P.
XX      14-FEB-2002; 2002US-0356119P.
XX      14-FEB-2002; 2002US-0356123P.
XX      14-FEB-2002; 2002US-0356126P.
XX      14-FEB-2002; 2002US-0356133P.
XX      14-FEB-2002; 2002US-0356135P.
XX      14-FEB-2002; 2002US-0356150P.
XX      14-FEB-2002; 2002US-0356151P.
XX      14-FEB-2002; 2002US-0356152P.
XX      14-FEB-2002; 2002US-0356154P.
XX      14-FEB-2002; 2002US-0356156P.
XX      14-FEB-2002; 2002US-0356157P.
XX      14-FEB-2002; 2002US-0356161P.
XX      14-FEB-2002; 2002US-0356162P.
XX
XX      (NOVA-) NOVAVAX INC.
XX
XX      Robinson RA;
XX
XX      WPI; 2003-646475/61.
XX      P-PSDB; ADA27363.
XX
XX      New insect cell line designated ATCC PTA-4047, useful for replicating
XX      baculoviruses to produce large amounts of recombinant proteins of
XX      medical, pharmaceutical and veterinary importance.
XX
XX      Claim 10; Fig 1; 63pp; English.
XX
XX      The present invention describes a cell line comprising a cell that is a
XX      clone, derivative, mutant and/or transfectant of a cell line designated
XX      American Type Culture Collection (ATCC) PTA-4047. The cell upon culture
XX      grows continuously and retains the identifying characteristics of the
XX      cell line designated ATCC-4047. Also described is a process of making a
XX      cell line. The insect cell line is useful in replicating baculoviruses,
XX      as a host substrate for baculovirus plaque assays, as a source of insect
XX      proteins, acts as a depot for cell transfection to produce recombinant
XX      baculoviruses, and in expressing viral recombinant proteins.
XX      Extracellular and intracellular viral recombinant proteins and virus-like
XX      particles expressed from the cell line are useful as pharmaceutical
XX      compositions, vaccines or diagnostic reagents. The present sequence
XX      represents a human papillomavirus type 16 (HPV-16) L1 codon optimised
XX      sequence, which is used in the exemplification of the present invention.
XX
XX      Sequence 1515 BP; 297 A; 532 C; 308 G; 278 T; 0 U; 0 Other;
XX
XX      Query Match      1.2%; Score 24; DB 8; Length 1515;
XX      Best Local Similarity 100.0%; Pred. No. 1;
XX      Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX      Qy      669 ACCTCTCCACCTCCACCGGCC 692
XX      |||||
XX      Db      1471 ACCTCTCCACCTCCACCGGCC 1494
XX
XX      RESULT 11
XX      ADA92549
XX      ID      ADA92549 standard; DNA; 1515 BP.
XX
XX      AC      ADA92549;
XX
XX      XX
XX      DT      20-NOV-2003 (first entry)
```

```
XX      HPV-16 L1 codon optimised DNA sequence SEQ ID NO:1.
XX
XX      codon optimised; viral capsid protein; virus-like particle; VLP;
XX      antigenic; human papillomavirus infection; virucide; vaccine;
XX      gene therapy; human papillomavirus type 16; dysplasia; infection; HPV-16;
XX      L1; gene; ds.
XX
XX      Synthetic.
XX      Human papillomavirus type 16.
XX
XX      Key      Location/Qualifiers
XX      CDS      1. .1515
XX              /*tag= a
XX              /product= "codon optimised L1 amino acid sequence"
XX              /transl_except= (pos:208. .210,aa:Tyr)
XX              /transl_except= (pos:295. .297,aa:Val)
XX
XX      WO2003068933-A2.
XX
XX      21-AUG-2003.
XX
XX      14-FEB-2003; 2003WO-US004480.
XX
XX      14-FEB-2002; 2002US-0356113P.
XX      14-FEB-2002; 2002US-0356118P.
XX      14-FEB-2002; 2002US-0356119P.
XX      14-FEB-2002; 2002US-0356123P.
XX      14-FEB-2002; 2002US-0356126P.
XX      14-FEB-2002; 2002US-0356133P.
XX      14-FEB-2002; 2002US-0356135P.
XX      14-FEB-2002; 2002US-0356150P.
XX      14-FEB-2002; 2002US-0356151P.
XX      14-FEB-2002; 2002US-0356152P.
XX      14-FEB-2002; 2002US-0356154P.
XX      14-FEB-2002; 2002US-0356156P.
XX      14-FEB-2002; 2002US-0356157P.
XX      14-FEB-2002; 2002US-0356161P.
XX      14-FEB-2002; 2002US-0356162P.
XX
XX      (NOVA-) NOVAVAX INC.
XX
XX      Robinson RA;
XX
XX      WPI; 2003-689664/65.
XX      P-PSDB; ADA92541.
XX
XX      New codon optimized polynucleotide encoding a viral capsid protein that
XX      self assembles into a virus-like particle, useful for diagnosing,
XX      preventing or treating human papillomavirus infections or associated
XX      disorders.
XX
XX      Claim 7; Page 113; 123pp; English.
XX
XX      The present invention describes a codon optimised polynucleotide encoding
XX      a viral capsid protein that self assembles into a virus-like particle
XX      (VLP) that exhibits conformational antigenic epitopes capable of raising
XX      neutralising antibodies, where the VLP is expressed from a host cell
XX      extracellularly. Also described: (1) a vector comprising the above codon
XX      optimised polynucleotide operably linked to a eukaryotic or prokaryotic
XX      regulatory control element, capable of replication in prokaryotic and/or
XX      eukaryotic host; (2) a host cell comprising the vector; (3) a
XX      pharmaceutical or vaccine composition for treating, ameliorating or
XX      preventing a papillomavirus related disease or disorder, comprising a
XX      multiplicity of VLPs that exhibit conformational antigenic epitopes, and
XX      a carrier, diluent or adjuvant; (4) a diagnostic kit for detecting a
XX      papillomavirus infection, comprising a multiplicity of VLPs that exhibit
XX      conformational antigenic epitopes, and a detection agent comprising a
XX      detectable label; (5) a method for preparing the above codon optimised
XX      polynucleotide, comprising replacing codons that are underutilised in
XX      insect cells with codons that are utilised at high levels in insect
XX      cells, to create an initially-modified nucleotide sequence, and modifying
XX      the initially-modified nucleotide sequence by choosing a preferred codon
```

CC for the initially-modified sequence, where the ratio of GC nucleotide
CC pairs to AT nucleotide pairs in the further-modified nucleotide sequence
CC trends towards about 1:1, where the number of palindromic and stem-loop
CC DNA structures in the further-modified nucleotide sequence is minimised,
CC and where the number of transcription and post-transcription repressor
CC elements are minimised; and (6) methods for treating, ameliorating or
CC preventing a papillomavirus related disease or disorder, or for
CC protecting an individual against a papillomavirus infection, comprising
CC administering to an individual an amount of the composition or vaccine
CC cited above. The VLP has virucide activity and can be used in vaccines
CC and in gene therapy. The composition and methods of the present invention
CC are useful in diagnosing, preventing or treating human papillomavirus
CC infections or associated disorders, such as dysplasia. The present
CC sequence represents an HPV-16 codon optimised L1 nucleotide sequence from
CC the present invention.

XX SQ Sequence 1515 BP; 297 A; 632 C; 308 G; 278 T; 0 U; 0 Other;
Query Match 1.2%; Score 24; DB 8; Length 1515;
Best Local Similarity 100.0%; Pred. No. 1;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 669 ACCTCTCCACCTCCACACCGCC 692
Db 1471 ACCTCTCCACCTCCACACCGCC 1494
|||||

RESULT 12
ADA14298
ID ADA14298 standard; DNA; 1515 BP.

XX AC ADA14298;

XX DT 06-NOV-2003 (first entry)

XX DE HPV-16 L1 codon optimised sequence SEQ ID NO:1.

XX purification; recombinant extracellular virus-like particle;
XX recombinant intracellular virus-like particle; virus-like particle; VLP;
XX virucide; vaccine; gene therapy; human papillomavirus; HPV; infection;
XX dysplasia; HPV-16; L1; codon optimised; gene; ds.

XX OS Human papillomavirus type 16.

XX FH Key Location/Qualifiers
XX CDS 1. .1515
FT /*tag= a
FT /product= "HPV-16 L1 protein"
FT /transl_except= (pos:208. .210,aa:Tyr)
FT /transl_except= (pos:295. .297,aa:Val)

XX PN WO2003068993-A1.

XX PD 21-AUG-2003.

XX PF 14-FEB-2003; 2003WO-US004474.

XX PR 14-FEB-2002; 2002US-0356113P.

XX PR 14-FEB-2002; 2002US-0356118P.

XX PR 14-FEB-2002; 2002US-0356119P.

XX PR 14-FEB-2002; 2002US-0356123P.

XX PR 14-FEB-2002; 2002US-0356126P.

XX PR 14-FEB-2002; 2002US-0356133P.

XX PR 14-FEB-2002; 2002US-0356135P.

XX PR 14-FEB-2002; 2002US-0356150P.

XX PR 14-FEB-2002; 2002US-0356151P.

XX PR 14-FEB-2002; 2002US-0356152P.

XX PR 14-FEB-2002; 2002US-0356154P.

XX PR 14-FEB-2002; 2002US-0356156P.

XX PR 14-FEB-2002; 2002US-0356157P.

XX PR 14-FEB-2002; 2002US-0356161P.

XX PR 14-FEB-2002; 2002US-0356162P.

XX

PA (NOVA-) NOVAVAX INC.

XX Robinson RA, Thompson MW;

XX WPI; 2003-679645/64.

DR P-PSDB; ADA14290.

XX

PT Purifying a recombinant human papillomavirus (HPV) L1, useful for
PT diagnosing, preventing or treating HPV infections, comprises clarifying,
PT concentrating and diafiltering cells containing HPV particles.

XX Example 4; Fig 1A-C; 111pp; English.

CC The present invention describes a method for purifying a recombinant
CC extracellular or intracellular virus-like particle (VLP). The method
CC comprises harvesting a cell suspension comprising cells containing a
CC plurality of VLPs to produce a harvested supernatant, optionally
CC disrupting the harvested cells to produce cell lysates containing the
CC VLP, clarifying the harvested supernatant, concentrating the clarified
CC supernatant, diafiltering the concentrated supernatant, and recovering
CC the purified recombinant VLP. Also described: (1) a cell line designated
CC as SF-9S deposited as American Type Culture Collection (ATCC) PTA-4047;
CC (2) producing the cell line described above; (3) host cells that express
CC one or more recombinant gene products with an enhanced yield; (4)
CC producing a foreign protein in an insect cell; (5) nucleic acid sequences
CC that correspond to and code for human papillomavirus (HPV) polypeptides;
CC and (6) pharmaceutical compositions comprising an amount of the
CC recombinant viral gene products, VLPs, agonists, antagonists, or the
CC active fragment of a viral gene product. The VLPs have virucide activity,
CC and can be used in vaccines and in gene therapy. The method is useful in
CC isolating and purifying expressed viral gene products, including VLPs, in
CC vitro. The gene products or particles may be used in detecting,
CC preventing or treating HPV infections and associated symptoms like
CC dysplasia. The present sequence represents an HPV-16 L1 codon optimised
CC nucleotide sequence, which is given in the exemplification of the present
CC invention.

XX SQ Sequence 1515 BP; 297 A; 632 C; 308 G; 278 T; 0 U; 0 Other;

Query Match 1.2%; Score 24; DB 8; Length 1515;

Best Local Similarity 100.0%; Pred. No. 1;

Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 669 ACCTCTCCACCTCCACACCGCC 692

Db 1471 ACCTCTCCACCTCCACACCGCC 1494
|||||

RESULT 13

AA058571

ID AA058571 standard; DNA; 1515 BP.

XX

XX AAD58571;

XX AC

XX 04-DEC-2003 (first entry)

XX DT

XX HPV-16 L1 codon optimised DNA.

XX DE

XX Vaccine; humoral immunity; cell-mediated immunity; gene therapy; HPV;
XX virucide; papillomavirus infection; gene; Human papillomavirus; ds.

XX OS

XX Human papillomavirus.

XX FH

XX Key Location/Qualifiers

XX CDS 1. .1515

XX /*tag= a

XX /product= "HPV-16 L1 protein"

XX FT

XX /transl_except= (pos:208. .210, aa:Tyr)

XX FT

XX /transl_except= (pos:295. .297, aa:Val)

XX FT

XX /note= "No stop codon"

XX FT

XX /partial

XX

XX WO2003068163-A2.

```
XX PD 21-AUG-2003.
XX PF 14-FEB-2003; 2003WO-US004473.
XX PR 14-FEB-2002; 2002US-0356113P.
XX PR 14-FEB-2002; 2002US-0356118P.
XX PR 14-FEB-2002; 2002US-0356119P.
XX PR 14-FEB-2002; 2002US-0356123P.
XX PR 14-FEB-2002; 2002US-0356126P.
XX PR 14-FEB-2002; 2002US-0356133P.
XX PR 14-FEB-2002; 2002US-0356135P.
XX PR 14-FEB-2002; 2002US-0356150P.
XX PR 14-FEB-2002; 2002US-0356151P.
XX PR 14-FEB-2002; 2002US-0356152P.
XX PR 14-FEB-2002; 2002US-0356154P.
XX PR 14-FEB-2002; 2002US-0356156P.
XX PR 14-FEB-2002; 2002US-0356157P.
XX PR 14-FEB-2002; 2002US-0356161P.
XX PR 14-FEB-2002; 2002US-0356162P.
XX PA (NOVA-) NOVAVAX INC.
XX PI Robinson RA, Cioce V;
XX PF WPI; 2003-689598/65.
XX DR P-PSDB; AAE38615.
XX PS Claim 10; Fig 1A-1C; 126pp; English.
XX CC The present invention relates to chimeric virus-like particle comprising
CC a recombinant viral capsid protein that encapsulates a recombinant viral
CC protein during self assembly into a chimeric virus-like particle and
CC exhibiting conformational antigenic epitopes capable of eliciting
CC neutralising antibodies. The vaccine comprising the chimeric virus-like
CC particles are useful for inducing immunity (humoral and/or cell-mediated
CC immunity) against papillomavirus infection. The invention is also useful
CC in gene therapy. The present sequence is HPV (human papillomavirus)-16 L1
CC codon optimised DNA
XX SQ Sequence 1515 BP; 297 A; 632 C; 308 G; 278 T; 0 U; 0 Other;
Query Match 1.2%; Score 24; DB 8; Length 1515;
Best Local Similarity 100.0%; Pred. No. 1;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 669 ACCTCTCCACCTCCACCCGCC 692
Db 1471 ACCTCTCCACCTCCACCCGCC 1494
RESULT 14
AAD47597
ID AAD47597 standard; DNA; 897 BP.
XX AC AAD47597;
XX DT 24-FEB-2003 (first entry)
XX DE Human HGFR M1250T mutant DNA fragment.
XX KW Human; hepatocyte growth factor receptor kinase; hHGFR; proto-oncogene;
XX protein co-ordinate data; drug design; hepatocyte growth factor; tumour;
XX therapy; enzyme; mutant; gene; ds.
XX OS Homo sapiens.
XX OS Synthetic.
XX FH Key Location/Qualifiers
1. .897
/*tag= a
/product= "Human HGFR mutant protein fragment"
```

```
FT CDS
1. .897
/*tag= a
/product= "Human HGFR mutant protein fragment"
/translation= (pos:424..426, aa:Ala)
/notes= "CDS does not include stop codon"
partial
FT PN EP1243596-A2.
XX PD 25-SEP-2002.
XX PF 25-MAR-2002; 2002EP-00006616.
XX PR 23-MAR-2001; 2001US-0277968P.
XX PA (AGOU-) AGOURON PHARM INC.
XX PI Mroczkowski B, Hickey M, Mctigue MA, Murray BW, Parge H, Sarup J;
XX PI Zhu J;
XX DR WPI; 2003-001650/01.
XX DR P-PSDB; AAE30037.
XX CC New polynucleotide encoding human hepatocyte growth factor receptor,
XX useful for designing modulators of the receptor, and potential antitumor
XX agents.
XX PS Claim 4; Page 157; 204pp; English.
XX CC The invention relates to human hepatocyte growth factor receptor kinase
XX (hHGFR) and its nucleic acid sequence. The invention also relates to
XX methods for identification of hHGFR inhibitors. hHGFR DNA is used to
XX express the corresponding polypeptides and these are used to provide
XX protein co-ordinate data useful in rational drug design of compounds that
XX modulate kinase domain, potentially useful for treating diseases
XX associated with signalling between hHGFR and hepatocyte growth factor,
XX particularly tumours (since HGFR is a proto-oncogene). It can also be
XX used to search databases for related sequences and its fragments as
XX primers or probes. The present sequence is human HGFR mutant DNA fragment
XX used to illustrate the method of the invention
XX SQ Sequence 897 BP; 257 A; 193 C; 207 G; 240 T; 0 U; 0 Other;
Query Match 1.2%; Score 23; DB 7; Length 897;
Best Local Similarity 100.0%; Pred. No. 3.1;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1240 AAGGTTGCTGATTTGCTTGC 1262
Db 508 AAGGTTGCTGATTTGCTTGC 530
RESULT 15
AAD47600
ID AAD47600 standard; DNA; 897 BP.
XX AC AAD47600;
XX DT 24-FEB-2003 (first entry)
XX DE Human HGFR H1094R mutant DNA fragment.
XX KW Human; hepatocyte growth factor receptor kinase; hHGFR; proto-oncogene;
XX protein co-ordinate data; drug design; hepatocyte growth factor; tumour;
XX therapy; enzyme; mutant; gene; ds.
XX OS Homo sapiens.
XX OS Synthetic.
XX FH Key Location/Qualifiers
1. .897
/*tag= a
/product= "Human HGFR mutant protein fragment"
```

FT /transl_except= (pos:424. .426, aa:Ala)
FT /transl_except= (pos:667. .669, aa:Leu)
FT /note= "CDS does not include stop codon"
FT /partial

XX EP1243596-A2.

XX PD 25-SEP-2002.

XX XX 25-MAR-2002; 2002EP-00006616.

XX XX 23-MAR-2001; 2001US-0277968P.

XX PA (AGOU-) AGOURON PHARM INC.

XX Mroczkowski B, Hickey M, Metigue MA, Murray BW, Parge H, Sarup J;
XX Zhu J;

DR WPI; 2003-001650/01.

DR P-PSDB; AAE30100.

XX New polynucleotide encoding human hepatocyte growth factor receptor,
PT useful for designing modulators of the receptor, and potential antitumor
PT agents.

XX PS Claim 6; Page 160-161; 204pp; English.

XX The invention relates to human hepatocyte growth factor receptor kinase
CC (hGFR) and its nucleic acid sequence. The invention also relates to
CC methods for identification of hGFR inhibitors. hGFR DNA is used to
CC express the corresponding polypeptides and these are used to provide
CC protein co-ordinate data useful in rational drug design of compounds that
CC modulate kinase domain, potentially useful for treating diseases
CC associated with signalling between hGFR and hepatocyte growth factor,
CC particularly tumours (since HGFR is a proto-oncogene). It can also be
CC used to search databases for related sequences and its fragments as
CC primers or probes. The present sequence is human HGFR mutant DNA fragment
CC used to illustrate the method of the invention

XX SQ Sequence 897 BP; 257 A; 191 C; 209 G; 240 T; 0 U; 0 Other;

Query Match 1.2%; Score 23; DB 7; Length 897;
Best Local Similarity 100.0%; Pred. No. 3.1;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1240 AAGTTGCTGATTTGGTCTTGC 1262

Db 508 AAGTTGCTGATTTGGTCTTGC 530

Search completed: May 12, 2004, 17:27:15
Job time : 750 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 12, 2004, 16:51:16 ; Search time 4545 Seconds
(without alignments)
12772.735 Million cell updates/sec

Title: US-10-086-464-1

Perfect score: 1944

Sequence: 1 atgctcctggcgccgtctcc.....atagtgacctctctttaa 1944

Scoring table: OLIGO_NUC

Gapop_60.0 , Gapext 60.0

Searched: 27513289 seqs, 14931090276 residues

Word size : 0

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmd:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_est1:*

10: gb_est2:*

11: gb_htc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vit:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_phg:*

27: em_gss_vrl:*

28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	429	22.1	723	14	CD839231
2	273	14.0	453	28	BH442220 BOGLN68TR
3	187	9.6	753	28	BZ450508 BONKA33TF
4	65	3.3	314	14	CD811645

5	65	3.3	314	14	CD811912	CD811912 BN10.019G
6	53	2.7	561	9	AV539333	AV539333 AV539333
7	53	2.7	584	9	AV543493	AV543493 AV543493
8	42	2.2	754	28	BH432371	BH432371 BOGVQ93TR
9	32	1.6	342	28	BZ456327	BZ456327 BONBB66TR
10	30	1.5	744	28	BZ066106	BZ066106 IJ170612
11	29	1.5	505	12	BM061076	BM061076 KSO1026B1
12	29	1.5	573	9	AV551753	AV551753 AV551753
13	29	1.5	607	29	CC968148	CC968148 BOIEB10TF
14	29	1.5	620	28	BH577077	BH577077 BOHCX77TR
15	29	1.5	651	13	BU043317	BU043317 PP_LEA001
16	29	1.5	695	28	BH928300	BH928300 Odi82F05
17	28	1.4	713	12	BP175991	BP175991 BP175991
18	27	1.4	414	28	BH434903	BH434903 BOGVQ04TF
19	27	1.4	536	9	AL369682	AL369682 MCB32F07
20	26	1.3	204	14	CD004164	CD004164 VVA031D03
21	26	1.3	295	14	F14384	F14384 ATTS5296 AC
22	26	1.3	316	9	AV547061	AV547061 AV547061
23	26	1.3	360	9	AV534188	AV534188 AV534188
24	26	1.3	395	9	AV819894	AV819894 AV819894
25	26	1.3	399	9	AV797870	AV797870 AV797870
26	26	1.3	402	9	AV800600	AV800600 AV800600
27	26	1.3	411	9	AV799812	AV799812 AV799812
28	26	1.3	417	9	AV810018	AV810018 AV810018
29	26	1.3	421	9	AV788292	AV788292 AV788292
30	26	1.3	430	9	AV795981	AV795981 AV795981
31	26	1.3	436	9	AV793009	AV793009 AV793009
32	26	1.3	439	9	AV819129	AV819129 AV819129
33	26	1.3	459	9	AV538962	AV538962 AV538962
34	26	1.3	482	13	BH209233	BH209233 UB16CPE10
35	26	1.3	484	12	BG817458	BG817458 EM1_76 CO
36	26	1.3	593	10	BP176907	BP176907 EM1_4_B10
37	26	1.3	595	14	CF015663	CF015663 OBLBb11.x
38	26	1.3	613	14	CF232052	CF232052 PfaJX0000
39	26	1.3	637	13	BQ134241	BQ134241 1091014D0
40	26	1.3	641	14	CA248316	CA248316 SCCCFU509
41	26	1.3	645	13	BQ506869	BQ506869 EST614284
42	26	1.3	666	13	BU499754	BU499754 946178G10
43	26	1.3	673	14	CA164704	CA164704 SCSBR2312
44	26	1.3	682	13	CA095337	CA095337 SCCCL500
45	26	1.3	750	13	CA148443	CA148443 SCUFR2100

ALIGNMENTS

RESULT 1:
CD839231
LOCUS RFO2.114I04F010529 RFO2 Brassica napus cDNA clone RFO2114I04, mRNA
DEFINITION RFO2.114I04F010529 RFO2 Brassica napus cDNA clone RFO2114I04, mRNA
sequence.
CD839231
VERSION CD839231.1 GI:32521171
KEYWORDS EST.
SOURCE Brassica napus (rape)
ORGANISM Brassica napus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE 1 (bases 1 to 723)
AUTHORS Genoplante.
TITLE Genoplante, a major partnership french program in plant genomics
JOURNAL Unpublished (2003)
COMMENT Contact: Genoplante
Genoplante
93, rue Henri Rochefort 91025 EVRY CEDEX France
Tel: 33 1 69 47 54 00
Fax: 33 1 69 47 54 10
This sequence has been generated in the framework of the french
plant genomics programme 'Genoplante' (<http://www.genoplante.com>)
and <http://genoplante-info.infobiogen.fr>.

FEATURES
1..723 source


```
/db_xref="taxon:3712"
/clone="BONKA33"
/clone_lib="BO_1.6.2_KB_tot"
/notes="Vector: PHOS1; Site: 1; BstXI; 1.6-2 kb sheared
total DNA inserted into PHOS1 using BstXI linkers"

ORIGIN
Query Match          9.6%; Score 187; DB 28; Length 753;
Best Local Similarity 100.0%; Pred. No. 2e-73;
Matches 187; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1282 AACACGATGATCAACACGCTGTGATGGAAACCTTTGGGTACTTGGCTCCGGAATACGCT 1341
Db |||||
Qy 11 AACACGATGATCAACACGCTGTGATGGAAACCTTTGGGTACTTGGCTCCGGAATACGCT 70
Db |||||
Qy 1342 GCACGCGAAGCTCACCGGAGCTGTGACCTTTCTCATTTGGCTGTCTTTGGAG 1401
Db |||||
Qy 71 GCAAGCGGAAGCTCACCGGAGCTGTGACCTTTCTCATTTGGCTGTCTTTGGAG 130
Db |||||
Qy 1402 CTCATTACTGACGCTGACCGCTTGATGCCAACATGTCTATGTAGATGACAGCTTAGTT 1461
Db |||||
Qy 131 CTCATTACTGACGCTGACCGCTTGATGCCAACATGTCTATGTAGATGACAGCTTAGTT 190
Db |||||
Qy 1462 GACTGGG 1468
Db |||||
Qy 191 GACTGGG 197
Db |||||

RESULT 4
CD811645
LOCUS          CD811645          314 bp      mRNA      linear      EST 10-JUL-2003
DEFINITION    BN10.001E08F011207 BN10 Brassica napus cDNA clone BN10001E08, mRNA
sequence.
ACCESSION    CD811645
VERSION      CD811645.1 GI:32493585
KEYWORDS     EST.
SOURCE       Brassica napus (rape)
ORGANISM     Brassica napus
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE    1 (bases 1 to 314)
AUTHORS      Genoplatne, a major partnership french program in plant genomics
TITLE        Unpublished (2003)
JOURNAL      Contact: Genoplatne
COMMENT      Genoplatne
93, rue Henri Rochefort 91025 EVRY CEDEX France
Tel: 33 1 69 47 54 00
Fax: 33 1 69 47 54 10
This sequence has been generated in the framework of the french
plant genomics programme 'Genoplatne' (http://www.genoplatne.com
and http://genoplatne-info.infobiogen.fr).

FEATURES             source
Location/Qualifiers
1..314
/organism="Brassica napus"
/mol_type="mRNA"
/cultivar="Jet neuf"
/db_xref="taxon:3708"
/clone="BN1001E08"
/tissue_type="seed"
/clone_lib="BN10"

ORIGIN
Query Match          3.3%; Score 65; DB 14; Length 314;
Best Local Similarity 99.1%; Pred. No. 3.8e-18;
Matches 115; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1825 CCGACCACTGACTGACTGACCTGTACCCGCTCTGTTCAAGCAGCGAGGGCCAAACACACGC 1884
Db |||||
Qy 1 CCGACCACTGACTGACTGACCTGTACCCGCTCTGTTCAAGCAGCGAGGGCCAAACACACGC 60
Db |||||
Qy 1885 GAAATGGAGATGGGGAAGATTAAAGAGAACCGGTCCAGGGTTATAGTGACCTTCTCT 1940
Db |||||
Qy 61 GAGATGGAGATGGGGAAGATTAAAGAGAACCGGTCCAGGGTTATAGTGACCTTCTCT 116
Db |||||

RESULT 6
AV539333/c
LOCUS          AV539333          561 bp      mRNA      linear      EST 07-SEP-2000
DEFINITION    AV539333 Arabidopsis thaliana roots Columbia Arabidopsis thaliana
cDNA clone RZ130C07F 3', mRNA sequence.
ACCESSION    AV539333
VERSION      AV539333.1 GI:8701090
KEYWORDS     EST.
SOURCE       Arabidopsis thaliana (thale cress)
ORGANISM     Arabidopsis thaliana
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE    1 (bases 1 to 561)
AUTHORS      Asamizu E., Nakamura Y., Sato S. and Tabata S.
TITLE        A large scale analysis of cDNA in Arabidopsis thaliana: Generation
of 12,028 non-redundant expressed sequence tags from normalized and
size-selected cDNA libraries
JOURNAL      DNA Res. 7 (3), 175-180 (2000)
MEDLINE      20363093
```

```
/db_xref="taxon:3712"
/clone="BONKA33"
/clone_lib="BO_1.6.2_KB_tot"
/notes="Vector: PHOS1; Site: 1; BstXI; 1.6-2 kb sheared
total DNA inserted into PHOS1 using BstXI linkers"

ORIGIN
Query Match          9.6%; Score 187; DB 28; Length 753;
Best Local Similarity 100.0%; Pred. No. 2e-73;
Matches 187; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1282 AACACGATGATCAACACGCTGTGATGGAAACCTTTGGGTACTTGGCTCCGGAATACGCT 1341
Db |||||
Qy 11 AACACGATGATCAACACGCTGTGATGGAAACCTTTGGGTACTTGGCTCCGGAATACGCT 70
Db |||||
Qy 1342 GCACGCGAAGCTCACCGGAGCTGTGACCTTTCTCATTTGGCTGTCTTTGGAG 1401
Db |||||
Qy 71 GCAAGCGGAAGCTCACCGGAGCTGTGACCTTTCTCATTTGGCTGTCTTTGGAG 130
Db |||||
Qy 1402 CTCATTACTGACGCTGACCGCTTGATGCCAACATGTCTATGTAGATGACAGCTTAGTT 1461
Db |||||
Qy 131 CTCATTACTGACGCTGACCGCTTGATGCCAACATGTCTATGTAGATGACAGCTTAGTT 190
Db |||||
Qy 1462 GACTGGG 1468
Db |||||
Qy 191 GACTGGG 197
Db |||||

RESULT 4
CD811645
LOCUS          CD811645          314 bp      mRNA      linear      EST 10-JUL-2003
DEFINITION    BN10.001E08F011207 BN10 Brassica napus cDNA clone BN10001E08, mRNA
sequence.
ACCESSION    CD811645
VERSION      CD811645.1 GI:32493585
KEYWORDS     EST.
SOURCE       Brassica napus (rape)
ORGANISM     Brassica napus
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE    1 (bases 1 to 314)
AUTHORS      Genoplatne, a major partnership french program in plant genomics
TITLE        Unpublished (2003)
JOURNAL      Contact: Genoplatne
COMMENT      Genoplatne
93, rue Henri Rochefort 91025 EVRY CEDEX France
Tel: 33 1 69 47 54 00
Fax: 33 1 69 47 54 10
This sequence has been generated in the framework of the french
plant genomics programme 'Genoplatne' (http://www.genoplatne.com
and http://genoplatne-info.infobiogen.fr).

FEATURES             source
Location/Qualifiers
1..314
/organism="Brassica napus"
/mol_type="mRNA"
/cultivar="Jet neuf"
/db_xref="taxon:3708"
/clone="BN1001E08"
/tissue_type="seed"
/clone_lib="BN10"

ORIGIN
Query Match          3.3%; Score 65; DB 14; Length 314;
Best Local Similarity 99.1%; Pred. No. 3.8e-18;
Matches 115; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1825 CCGACCACTGACTGACTGACCTGTACCCGCTCTGTTCAAGCAGCGAGGGCCAAACACACGC 1884
Db |||||
Qy 1 CCGACCACTGACTGACTGACCTGTACCCGCTCTGTTCAAGCAGCGAGGGCCAAACACACGC 60
Db |||||
Qy 1885 GAAATGGAGATGGGGAAGATTAAAGAGAACCGGTCCAGGGTTATAGTGACCTTCTCT 1940
Db |||||
Qy 61 GAGATGGAGATGGGGAAGATTAAAGAGAACCGGTCCAGGGTTATAGTGACCTTCTCT 116
Db |||||

RESULT 6
AV539333/c
LOCUS          AV539333          561 bp      mRNA      linear      EST 07-SEP-2000
DEFINITION    AV539333 Arabidopsis thaliana roots Columbia Arabidopsis thaliana
cDNA clone RZ130C07F 3', mRNA sequence.
ACCESSION    AV539333
VERSION      AV539333.1 GI:8701090
KEYWORDS     EST.
SOURCE       Arabidopsis thaliana (thale cress)
ORGANISM     Arabidopsis thaliana
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE    1 (bases 1 to 561)
AUTHORS      Asamizu E., Nakamura Y., Sato S. and Tabata S.
TITLE        A large scale analysis of cDNA in Arabidopsis thaliana: Generation
of 12,028 non-redundant expressed sequence tags from normalized and
size-selected cDNA libraries
JOURNAL      DNA Res. 7 (3), 175-180 (2000)
MEDLINE      20363093
```

PUBMED 10907847
COMMENT Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.
FEATURES
source
1..561
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strain="Columbia"
/db_xref="taxon:3702"
/clone="RZ130C07F"
/tissue_type="roots"
/clone_lib="Arabidopsis thaliana roots Columbia"
/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2: XhoI"
ORIGIN
Query Match 2.7%; Score 53; DB 9; Length 561;
Best Local Similarity 100.0%; Pred. No. 1.2e-12;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1579 GCTCGGCTTGTTGGCATTTCAGCTCGCGCAGACCTCGCATGAGCCAGAT 1631
DB 543 GCTCGGCTTGTTGGCATTTCAGCTCGCGCAGACCTCGCATGAGCCAGAT 491
RESULT 7
AV543493/c
LOCUS AV543493 Arabidopsis thaliana roots Columbia Arabidopsis thaliana
DEFINITION CDNA clone RZ201f07F 3', mRNA sequence.
ACCESSION AV543493.1 GI:8714907
VERSION EST.
KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE 1 (bases 1 to 584)
AUTHORS Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S.
TITLE A large scale analysis of cDNA in Arabidopsis thaliana: Generation of 12,028 non-redundant expressed sequence tags from normalized and size-selected cDNA libraries
JOURNAL DNA Res. 7 (3), 175-180 (2000)
MEDLINE 20363093
PUBMED 10907847
COMMENT Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.
FEATURES
source
1..584
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strain="Columbia"
/db_xref="taxon:3702"
/clone="RZ201f07F"
/tissue_type="roots"
/clone_lib="Arabidopsis thaliana roots Columbia"
/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2: XhoI"
ORIGIN
Query Match 2.7%; Score 53; DB 9; Length 584;
Best Local Similarity 100.0%; Pred. No. 1.3e-12;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1579 GCTCGGCTTGTTGGCATTTCAGCTCGCGCAGACCTCGCATGAGCCAGAT 1631
|||||

DB 468 GCTCGGCTTGTTGGCATTTCAGCTCGCGCAGACCTCGCATGAGCCAGAT 416
RESULT 8
BH432371
LOCUS BH432371 754 bp DNA linear GSS 12-DEC-2001
DEFINITION BOGVQ93TR BOGV Brassica oleracea genomic clone BOGVQ93, genomic survey sequence.
ACCESSION BH432371.1 GI:17618092
VERSION BH432371
KEYWORDS GSS.
SOURCE Brassica oleracea
ORGANISM Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE 1 (bases 1 to 754)
AUTHORS Town, C.D., Van Aken, S., Utterback, T., Koo, H. and Fraser, C.M.
TITLE Whole genome shotgun sequencing of Brassica oleracea
JOURNAL Unpublished (2001)
COMMENT Other GSSs: BOGVQ93TF
Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TR
Class: sheared ends.
FEATURES
source
1..754
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/strain="TO1000DH3"
/db_xref="taxon:3712"
/clone="BOGVQ93"
/clone_lib="BOGV"
/note="Vector: pHOS1; Site_1: BstXI; 2-3 kb sheared genomic DNA inserted into pHOS1 using BstXI linkers"
ORIGIN
Query Match 2.2%; Score 42; DB 28; Length 754;
Best Local Similarity 100.0%; Pred. No. 1.3e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1762 AAGAAGTTTAGGAAAATGGCACTTGGAACTCAAGAGTACAAC 1803
|||||
DB 462 AAGAAGTTTAGGAAAATGGCACTTGGAACTCAAGAGTACAAC 503
RESULT 9
BZ456327
LOCUS BZ456327 BO_1.6_2_KB_tot Brassica oleracea genomic clone BONBB66,
DEFINITION genomic survey sequence.
ACCESSION BZ456327
VERSION BZ456327.1 GI:26734454
KEYWORDS GSS.
SOURCE Brassica oleracea
ORGANISM Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE 1 (bases 1 to 342)
AUTHORS Town, C.D., Van Aken, S., Utterback, T., Koo, H. and Fraser, C.M.
TITLE Whole genome shotgun sequencing of Brassica oleracea
JOURNAL Unpublished (2001)
COMMENT Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208

Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TR
Class: sheared ends.

FEATURES

source Location/Qualifiers

1..342
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/strain="TO1000DH3"
/db_xref="taxon:3712"
/clone="BONBB66"
/clone_lib="BO 1.6.2 KB tot"
/note="Vector: pHS1; Site 1: BstXI; 1.6-2 kb sheared
total DNA inserted into pHS1 using BstXI linkers"

ORIGIN

Query Match 1.6%; Score 32; DB 28; Length 342;
Best Local Similarity 100.0%; Pred. No. 0.0038;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 37 CCATCTCCACCATCAAACTCCACACACACAC 68

Db 182 CCATCTCCACCATCAAACTCCACACACAC 213

RESULT 10

BZ066106/c

LOCUS

DEFINITION 13170el2.g1 B.oleracea002 Brassica oleracea genomic, genomic survey
sequence.

ACCESSION BZ066106

VERSION BZ066106.1

KEYWORDS GSS.

SOURCE Brassica oleracea

ORGANISM Brassica oleracea

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

REFERENCE 1 (bases 1 to 744)

AUTHORS Dalehaunty,K., Fellw.G., Fulton,L., McCombie,W.R., Miner,T.,

Nash,W., Rabinowicz,P.D. and Wilson,R.K.

TITLE Whole genome shotgun reads from Brassica oleracea

JOURNAL Unpublished (2002)

COMMENT Contact: Richard K. Wilson

Genome Sequencing Center

Washington University School of Medicine

Email: submissions@watson.wustl.edu

Plate: 13170 row: e column: 12

Seq primer: -28RppOT reverse

Class: shotgun

High quality sequence start: 17

High quality sequence stop: 551.

FEATURES

source

1..744
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/db_xref="taxon:3712"
/clone_lib="B.oleracea002"
/note="Vector: pOTw13; Whole genome shotgun library from
flowering buds. DNA was purified from a crude nuclear
prep using Brassica oleracea TO1000DH3 buds provided by
Thomas Osborn at the University of Wisconsin. Genomic
DNA was provided by Pablo Rabinowicz (CSHL) and the
shotgun library prepared at Washington University Genome
Sequencing Center."

ORIGIN

Query Match 1.5%; Score 30; DB 28; Length 744;
Best Local Similarity 100.0%; Pred. No. 0.037;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1596 CCATTCAGCTCGCGCAGACCTCGCATGAG 1625

|||||

Db 729 CCATTCAGCTCGCGCAGACCTCGCATGAG 700

RESULT 11

BM061076

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

BM061076

KS01026B11

KS01

Capsicum annuum

CDNA, mRNA

sequence.

EST.

BM061076.1

GI:22781194

EST.

Capsicum annuum

Capsicum annuum

Capsicum annuum

Capsicum annuum

Capsicum annuum

Capsicum annuum

Capsicum annuum

Capsicum annuum

Capsicum annuum

Capsicum annuum

Capsicum annuum

Capsicum annuum

Capsicum annuum

Capsicum annuum

Capsicum annuum

Capsicum annuum

Capsicum annuum

Capsicum annuum

Capsicum annuum

Capsicum annuum

Capsicum annuum

Capsicum annuum

Capsicum annuum

Capsicum annuum

Capsicum annuum

Capsicum annuum

Capsicum annuum

Capsicum annuum

Capsicum annuum

Capsicum annuum

Capsicum annuum

Capsicum annuum

Capsicum annuum

Capsicum annuum

Capsicum annuum

Capsicum annuum

Capsicum annuum

Capsicum annuum

Capsicum annuum

Capsicum annuum

Capsicum annuum

Capsicum annuum

Capsicum annuum

Capsicum annuum

Capsicum annuum

Capsicum annuum

Capsicum annuum

Capsicum annuum

Capsicum annuum

Capsicum annuum

Capsicum annuum

Capsicum annuum

Capsicum annuum

```

COMMENT      Contact: Erika Asamizu
              The First Laboratory for Plant Gene Research
              Kazusa DNA Research Institute
              Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
              Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.

FEATURES
  source
    1..573
      /organism="Arabidopsis thaliana"
      /mol_type="mRNA"
      /strain="Columbia"
      /db_xref="taxon:3702"
      /clone="R2130c07p"
      /tissue_type="roots"
      /clone_lib="Arabidopsis thaliana roots Columbia"
      /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
      XhoI"

ORIGIN
Query Match      1.5%; Score 29; DB 9; Length 573;
Best Local Similarity 100.0%; Pred. No. 0.099;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      880  TTGCTAGTGGGAAGAGTTGCTGTGAA 908
          |||
          471  TTGCTAGTGGGAAGAGTTGCTGTGAA 499

RESULT 13
CC968148      607 bp      DNA      linear      GSS 18-AUG-2003
LOCUS      BOIEB10TF BO_1.4_1.6_KB_nuc Brassica oleracea genomic clone
DEFINITION      BOIEB10, genomic survey sequence.
ACCESSION      CC968148
VERSION        CC968148.1 GI:33822075
KEYWORDS       GSS.
SOURCE         Brassica oleracea
ORGANISM       Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE      1 (bases 1 to 607)
AUTHORS        Town,C.D., Van Aken,S., Utterback,T., Koo,H. and Fraser,C.M.
TITLE          Whole genome shotgun sequencing of Brassica oleracea
JOURNAL        Unpublished (2001)
COMMENT        Other GSSs: BOIEB10TR
                Contact: Chris Town
                TIGR
                9712 Medical Center Drive, Rockville, MD 20850, USA.
                Tel: 301-838-3523
                Fax: 301-838-0208
                Email: cdtown@tigr.org
                DNA is from a doubled haploid provided by Tom Osborn.
                Seq primer: TF
                Class: sheared ends.
                Location/Qualifiers
                  1..607
                  /organism="Brassica oleracea"
                  /mol_type="genomic DNA"
                  /strain="TO1000DH3"
                  /db_xref="taxon:3712"
                  /clone="BOIEB10"
                  /clone_lib="BOIEB10"
                  /note="Vector: pHOS2; Site 1: BstXI; 1.4-1.6 kb sheared
                  nuclear DNA inserted into pHOS2 using BstXI linkers"

FEATURES
  source
    1..607
      /organism="Brassica oleracea"
      /mol_type="genomic DNA"
      /strain="TO1000DH3"
      /db_xref="taxon:3712"
      /clone="BOIEB10"
      /clone_lib="BOIEB10"
      /note="Vector: pHOS2; Site 1: BstXI; 1.4-1.6 kb sheared
      nuclear DNA inserted into pHOS2 using BstXI linkers"

ORIGIN
Query Match      1.5%; Score 29; DB 29; Length 607;
Best Local Similarity 100.0%; Pred. No. 0.1;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      871  AAAGTGTTGCTAGTGGGAAGAGT 899
          |||
          431  AAAGTGTTGCTAGTGGGAAGAGT 459

RESULT 15
CC968148      651 bp      mRNA      linear      EST 26-AUG-2002
LOCUS      PP LEA0015K13f Peach developing fruit mesocarp Prunus persica cDNA
DEFINITION      BU043317
ACCESSION      BU043317
VERSION        BU043317.1 GI:22483394
KEYWORDS       EST.
SOURCE         Prunus persica (peach)
ORGANISM       Prunus persica
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Rosales; Rosaceae; Amygdaloideae; Prunus.
REFERENCE      1 (bases 1 to 651)
AUTHORS        Callahan,A., Palmer,M., Main,D., Wing,R. and Abbott,A.
TITLE          Peach Model Genome for Rosaceae
JOURNAL        Unpublished (2002)
COMMENT        Contact: Abbott, A.
                Dept of Genetics and Biochemistry
                Clemson University
                122 Long Hall, Clemson University, Clemson, SC 29634, USA
                Tel: 864 656 3060
                Fax: 864 656 6879

```

```

RESULT 14
BH577077/c
LOCUS      BH577077
DEFINITION      BH577077 BOHC Brassica oleracea genomic clone BOHCX77, genomic
                survey sequence.
ACCESSION      BH577077
VERSION        BH577077.1 GI:17829358
KEYWORDS       GSS.
SOURCE         Brassica oleracea
ORGANISM       Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE      1 (bases 1 to 620)
AUTHORS        Town,C.D., Van Aken,S., Utterback,T., Koo,H. and Fraser,C.M.
TITLE          Whole genome shotgun sequencing of Brassica oleracea
JOURNAL        Unpublished (2001)
COMMENT        Other GSSs: BOHCX77TF
                Contact: Chris Town
                TIGR
                9712 Medical Center Drive, Rockville, MD 20850, USA.
                Tel: 301-838-3523
                Fax: 301-838-0208
                Email: cdtown@tigr.org
                DNA is from a doubled haploid provided by Tom Osborn.
                Seq primer: TR
                Class: sheared ends.
                Location/Qualifiers
                  1..620
                  /organism="Brassica oleracea"
                  /mol_type="genomic DNA"
                  /strain="TO1000DH3"
                  /db_xref="taxon:3712"
                  /clone="BOHCX77"
                  /clone_lib="BOHC"
                  /note="Vector: pHOS1; Site 1: BstXI; 2-3 kb sheared
                  genomic DNA inserted into pHOS1 using BstXI linkers"

ORIGIN
Query Match      1.5%; Score 29; DB 28; Length 620;
Best Local Similarity 100.0%; Pred. No. 0.1;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      871  AAAGTGTTGCTAGTGGGAAGAGT 899
          |||
          567  AAAGTGTTGCTAGTGGGAAGAGT 539

RESULT 15
BU043317
LOCUS      PP LEA0015K13f Peach developing fruit mesocarp Prunus persica cDNA
DEFINITION      BU043317
ACCESSION      BU043317
VERSION        BU043317.1 GI:22483394
KEYWORDS       EST.
SOURCE         Prunus persica (peach)
ORGANISM       Prunus persica
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Rosales; Rosaceae; Amygdaloideae; Prunus.
REFERENCE      1 (bases 1 to 651)
AUTHORS        Callahan,A., Palmer,M., Main,D., Wing,R. and Abbott,A.
TITLE          Peach Model Genome for Rosaceae
JOURNAL        Unpublished (2002)
COMMENT        Contact: Abbott, A.
                Dept of Genetics and Biochemistry
                Clemson University
                122 Long Hall, Clemson University, Clemson, SC 29634, USA
                Tel: 864 656 3060
                Fax: 864 656 6879

```

Email: aalbert@clemson.edu
Total High Quality bases = 578
Seq primer: TAATACGACTCACTATAGG
High quality sequence stop: 651.

FEATURES

Source

Location/Qualifiers
1..651
/organism="Prunus persica"
/mol_type="mRNA"
/cultivar="Loring"
/db_xref="taxon:3760"
/clone="PP_LEA0015K13f"
/tissue_type="Mesocarp"
/lab_host="E. coli"
/clone_lib="Peach developing fruit mesocarp"
/note="Vector: pBluescript II SK(-); Site 1: EcoRI;
Site 2: XhoI; authority=Prunus persica L. Batsh; The
sequence has been trimmed to remove vector sequence and
contains a minimum of 100 bases of phred value 20 or
above. For more details on library preparation and
sequence analysis go to
<http://www.genome.clemson.edu/projects/peach>. To order
this clone go to <http://www.genome.clemson.edu/orders>"

ORIGIN

Query Match 1.5%; Score 29; DB 13; Length 651;
Best Local Similarity 100.0%; Pred.No. 0.1;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1555 GAGATGGCTCGCATGGTTGCTTGCTGCTGC 1583
|||
Db 270 GAGATGGCTCGCATGGTTGCTTGCTGCTGC 298

Search completed: May 12, 2004, 20:46:18
Job time : 4562 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 12, 2004, 16:54:59 ; Search time 168 Seconds
(without alignments)
6421.575 Million cell updates/sec

Title: US-10-086-464-1
Perfect score: 1944
Sequence: 1 atgctcctggcgccgtctcc.....atagtgacattctcttaa 1944

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 682709 seqs, 277475446 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents_NA.*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	23	1.2	4626	US-08-306-691B-22	Sequence 22, Appl
2	23	1.2	4626	PCT-US93-06251-27	Sequence 27, Appl
3	20	1.0	609	US-09-800-729-55	Sequence 55, Appl
4	20	1.0	852	US-09-252-991A-1274	Sequence 1274, Ap
5	20	1.0	909	US-09-252-991A-1381	Sequence 1381, Ap
6	20	1.0	1092	US-09-252-991A-1287	Sequence 1287, Ap
7	20	1.0	3097	US-09-228-986-11	Sequence 11, Appl
8	20	1.0	5720	US-09-800-729-18	Sequence 18, Appl
9	20	1.0	41100	US-09-755-665-46	Sequence 46, Appl
10	19	1.0	358	US-08-991-789A-168	Sequence 168, App
11	19	1.0	358	US-09-062-451-168	Sequence 168, App
12	19	1.0	358	US-09-598-326-168	Sequence 168, App
13	19	1.0	358	US-09-289-198-168	Sequence 168, App
14	19	1.0	358	US-09-429-755-168	Sequence 168, App
15	19	1.0	2421	US-09-187-330-4	Sequence 4, Appl
16	19	1.0	2487	US-09-187-330-54	Sequence 54, Appl
17	19	1.0	3621	US-09-635-872A-21	Sequence 21, Appl
18	19	1.0	3621	US-09-636-077A-21	Sequence 21, Appl
19	19	1.0	3621	US-09-636-060C-21	Sequence 21, Appl
20	19	1.0	3621	US-09-986-552-21	Sequence 21, Appl
21	19	1.0	3783	US-09-635-872A-20	Sequence 20, Appl
22	19	1.0	3783	US-09-636-077A-20	Sequence 20, Appl
23	19	1.0	3783	US-09-636-060C-20	Sequence 20, Appl
24	19	1.0	3783	US-09-986-552-20	Sequence 20, Appl
25	19	1.0	5053	US-09-376-330-1	Sequence 1, Appl
26	19	1.0	5597	US-09-635-872A-4	Sequence 4, Appl
27	19	1.0	5597	US-09-636-077A-4	Sequence 4, Appl

c 28	19	1.0	5597	4	US-09-636-060C-4	Sequence 4, Appli
c 29	19	1.0	5597	4	US-09-986-552-4	Sequence 4, Appli
c 30	18	0.9	274	4	US-09-313-294A-2260	Sequence 2260, Ap
c 31	18	0.9	293	4	US-09-313-294A-6390	Sequence 6390, Ap
c 32	18	0.9	300	4	US-09-313-294A-4708	Sequence 4708, Ap
c 33	18	0.9	306	4	US-09-313-294A-4079	Sequence 4079, Ap
c 34	18	0.9	308	4	US-09-313-294A-5000	Sequence 5000, Ap
c 35	18	0.9	424	4	US-09-621-976-9270	Sequence 9270, Ap
c 36	18	0.9	439	4	US-09-397-787-273	Sequence 273, App
c 37	18	0.9	489	1	US-07-879-685B-3	Sequence 3, Appli
c 38	18	0.9	652	4	US-09-976-594-666	Sequence 666, App
c 39	18	0.9	657	4	US-09-833-381-521	Sequence 921, App
c 40	18	0.9	684	3	US-08-998-416-672	Sequence 672, App
c 41	18	0.9	1008	4	US-09-500-569-1	Sequence 1, Appli
c 42	18	0.9	1008	4	US-09-971-823B-1	Sequence 1, Appli
c 43	18	0.9	1276	3	US-09-177-325-2	Sequence 2, Appli
c 44	18	0.9	1276	3	US-09-411-812A-2	Sequence 2, Appli
c 45	18	0.9	1276	4	US-09-590-113-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-08-306-691B-22
; Sequence 22, Application US/08306691B
; Patent No. 5734039
; GENERAL INFORMATION:
; APPLICANT: Calabretta, Bruno
; APPLICANT: Skorski, Tomasz
; TITLE OF INVENTION: OLIGONUCLEOTIDES TARGETING COOPERATING ONCOGENES
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Seidel, Gonda, Lavorgha & Monaco, P.C.
; STREET: Two Penn Center, Suite 1800
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 720 KB
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/306,691B
; FILING DATE: September 15, 1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 8321-8
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; TELEX: No. 5734039e
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4626 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-306-691B-22

Query Match 1-2%; Score 23; DB 1; Length 4626;
Best Local Similarity 100.0%; Pred. No. 0.27; Indels 0; Gaps 0;
Matches 23; Conservative 0; Mismatches 0;
Oy 1240 AAGGTGCTGATTGTTGCTTGC 1262

Qy 1467 GGCACGACCATGCTTAACC 1486
Db 112 GGCACGACCATGCTTAACC 93

RESULT 6

US-09-252-991A-1287/c
; Sequence 1287, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 1287
; LENGTH: 1092
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-1287

Query Match 1.0%; Score 20; DB 4; Length 1092;
Best Local Similarity 100.0%; Pred. No. 7.4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1467 GGCACGACCATGCTTAACC 1486
Db 126 GGCACGACCATGCTTAACC 107

RESULT 7

US-09-228-986-11
; Sequence 11, Application US/09228986
; Patent No. 6359198
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Niels
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signalling
; FILE REFERENCE: 11000/1020
; CURRENT APPLICATION NUMBER: US/09/228,986
; CURRENT FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 130
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 3097
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-09-228-986-11

Query Match 1.0%; Score 20; DB 4; Length 3097;
Best Local Similarity 100.0%; Pred. No. 7.5;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1387 GTTGCTCTTTGGAGCTCAT 1406
Db 2556 GTTGCTCTTTGGAGCTCAT 2575

RESULT 8

US-09-800-729-18
; Sequence 18, Application US/09800729
; Patent No. 6605592
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: 32 Human secreted proteins
; FILE REFERENCE: P2044P1

; CURRENT APPLICATION NUMBER: US/09/800,729
; CURRENT FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: PCT/US00/26013
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 60/155,709
; PRIOR FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 5720
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-800-729-18

Query Match 1.0%; Score 20; DB 4; Length 5720;
Best Local Similarity 100.0%; Pred. No. 7.5;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 667 CAACCTCTCCACCTCCACC 686
Db 5431 CAACCTCTCCACCTCCACC 5450

RESULT 9

US-09-755-665-46/c
; Sequence 46, Application US/09755665
; Patent No. 6600019
; GENERAL INFORMATION:
; APPLICANT: Prayaga, Sudhirdas K.
; APPLICANT: Majumder, Kumud
; APPLICANT: Tailon, Bruce E.
; APPLICANT: Spaderna, Steven K.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: MacDougall, John
; TITLE OF INVENTION: NOVEL POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-631
; CURRENT APPLICATION NUMBER: US/09/755,665
; CURRENT FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/174,724
; PRIOR FILING DATE: 2000-01-06
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 46
; LENGTH: 41100
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: GENOMIC DNA
US-09-755-665-46

Query Match 1.0%; Score 20; DB 4; Length 41100;
Best Local Similarity 100.0%; Pred. No. 7.5;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 921 TGGGAGTGGTCAGGAGAGA 940
Db 22048 TGGGAGTGGTCAGGAGAGA 22029

RESULT 10

US-08-991-789A-168
; Sequence 168, Application US/08991789A
; Patent No. 6225054
; GENERAL INFORMATION:
; APPLICANT: Fridakis, Tony N.
; APPLICANT: Smith, John M.
; APPLICANT: Reed, Steven G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TREATMENT AND DIAGNOSIS OF BREAST CANCER
; NUMBER OF SEQUENCES: 292
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed IP Law Group
; STREET: 701 Fifth Avenue, Suite 6300

```
;
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
;
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/991,789A
; FILING DATE: 11-Dec-1997
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Potter, Jane E. R.
; REGISTRATION NUMBER: 33,332
; REFERENCE/DOCKET NUMBER: 210121.419C3
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 168:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 358 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 168:
US-08-991-789A-168

Query Match 1.0%; Score 19; DB 3; Length 358;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 950 AGGCAGAGGTTGAGATCAT 968
Db 228 AGGCAGAGGTTGAGATCAT 246

RESULT 11
US-09-062-451-168
; Sequence 168, Application US/09062451
; Patent No. 6344550
; GENERAL INFORMATION:
; APPLICANT: Frudakis, Tony N.
; APPLICANT: Smith, John M.
; APPLICANT: Reed, Steven G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF BREAST CANCER
; NUMBER OF SEQUENCES: 297
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/062,451
; FILING DATE: 04-APR-1997
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.419C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 168:
```

```
;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 358 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-062-451-168

Query Match 1.0%; Score 19; DB 4; Length 358;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 950 AGGCAGAGGTTGAGATCAT 968
Db 228 AGGCAGAGGTTGAGATCAT 246

RESULT 12
US-09-598-326-168
; Sequence 168, Application US/09598326
; Patent No. 6423496
; GENERAL INFORMATION:
; APPLICANT: Frudakis, Tony N.
; APPLICANT: Smith, John M.
; APPLICANT: Reed, Steven G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF BREAST CANCER
; NUMBER OF SEQUENCES: 247
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed Intellectual Property Law Group PLLC
; STREET: 701 Fifth Avenue, Suite 6300
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/598,326
; FILING DATE: 20-Jun-2000
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Potter, Jane E.R.
; REGISTRATION NUMBER: 33,332
; REFERENCE/DOCKET NUMBER: 210121.419D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 168:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 358 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 168:
US-09-598-326-168

Query Match 1.0%; Score 19; DB 4; Length 358;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 950 AGGCAGAGGTTGAGATCAT 968
Db 228 AGGCAGAGGTTGAGATCAT 246

RESULT 13
US-09-289-198-168
; Sequence 168, Application US/09289198
; Patent No. 6586570
; GENERAL INFORMATION:
```

; APPLICANT: Frudakis, Tony N.
; APPLICANT: Smith, John M.
; APPLICANT: Reed, Steven G.
; APPLICANT: Misher, Lynda
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.419C5
; CURRENT APPLICATION NUMBER: US/09/289,198
; CURRENT FILING DATE: 1999-04-09
; EARLIER APPLICATION NUMBER: US 09/062,451
; EARLIER FILING DATE: 1998-04-17
; EARLIER APPLICATION NUMBER: US 08/991,789
; EARLIER FILING DATE: 1997-12-11
; EARLIER APPLICATION NUMBER: US 08/838,762
; EARLIER FILING DATE: 1997-04-09
; EARLIER APPLICATION NUMBER: PCT/US97/00485
; EARLIER FILING DATE: 1997-01-10
; EARLIER APPLICATION NUMBER: US 08/700,014
; EARLIER FILING DATE: 1996-08-20
; EARLIER APPLICATION NUMBER: US 08/585,392
; EARLIER FILING DATE: 1996-01-01
; NUMBER OF SEQ ID NOS: 312
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 168
; LENGTH: 358
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-289-198-168

Query Match 1.0%; Score 19; DB 4; Length 358;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 950 AGGCAGAGGTTGAGATCAT 968
Db 228 AGGCAGAGGTTGAGATCAT 246

RESULT 14

US-09-429-755-168

; Sequence 168, Application US/09429755A
; Patent No. 6656480
; GENERAL INFORMATION:
; APPLICANT: Frudakis, Tony N.
; APPLICANT: Smith, John M.
; APPLICANT: Reed, Steven G.
; APPLICANT: Misher, Lynda
; APPLICANT: Retter, Marc W.
; APPLICANT: Dillon, Davin C.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.419C6
; CURRENT APPLICATION NUMBER: US/09/429,755A
; CURRENT FILING DATE: 1999-10-28
; NUMBER OF SEQ ID NOS: 315
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 168
; LENGTH: 358
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-429-755-168

Query Match 1.0%; Score 19; DB 4; Length 358;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 950 AGGCAGAGGTTGAGATCAT 968
Db 228 AGGCAGAGGTTGAGATCAT 246

RESULT 15

US-09-187-330-4/c

; Sequence 4, Application US/09187330
; Patent No. 6613740
; GENERAL INFORMATION:
; APPLICANT: Gozes, Illana
; APPLICANT: Brennenman, Douglas E.
; APPLICANT: Bassan, Merav
; APPLICANT: Zamostiano, Rachel
; APPLICANT: The Government of the United States of America
; APPLICANT: as represented by the Secretary of the
; APPLICANT: Department of Health and Human Services
; TITLE OF INVENTION: Activity Dependent Neurotrophic Factor III (ADNF III)
; FILE REFERENCE: 015280-291200US
; CURRENT APPLICATION NUMBER: US/09/187,330
; CURRENT FILING DATE: 1998-11-06
; EARLIER APPLICATION NUMBER: US 60/037,404
; EARLIER FILING DATE: 1997-02-07
; EARLIER APPLICATION NUMBER: WO PCT/US98/02485
; EARLIER FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 2421
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; OTHER INFORMATION: one mouse activity dependent neurotrophic factor
; OTHER INFORMATION: III (ADNF III) cDNA clone
US-09-187-330-4

Query Match 1.0%; Score 19; DB 4; Length 2421;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 69 TCCTCCTCCAGCTTCGCGCT 87
Db 1295 TCCTCCTCCAGCTTCGCGCT 1277

Search completed: May 12, 2004, 20:49:09

Job time : 173 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 12, 2004, 19:30:15 ; Search time 837 Seconds
(without alignments)
10521.270 Million cell updates/sec

Title: US-10-086-464-1

Perfect score: 1944

Sequence: 1 atgtcctcgccgctctcc.....atagtggaccttctttaa 1944

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 2941586 seqs, 2264995651 residues

Word size : 0

Total number of hits satisfying chosen parameters: 5883172

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1944	100.0	1944	14	US-10-086-464-1
2	1944	100.0	2189	14	US-10-086-464-3
3	26	1.3	268	9	US-09-923-876-5096
4	26	1.3	268	11	US-09-923-876-5096
5	26	1.3	951	13	US-10-425-114-1922
6	26	1.3	1011	13	US-10-425-114-13131
7	26	1.3	1134	13	US-10-425-114-36420
8	26	1.3	1500	13	US-10-425-114-6300
9	26	1.3	1552	13	US-10-424-599-47702
10	26	1.3	1724	13	US-10-425-114-2326
11	25	1.3	1101	13	US-10-424-599-122411
12	24	1.2	1515	13	US-10-368-046-1
13	24	1.2	1515	16	US-10-367-095-1
14	23	1.2	180	15	US-10-029-386-26494

15	23	1.2	579	15	US-10-029-386-12794	Sequence 12794, A
16	23	1.2	897	16	US-10-106-938A-10	Sequence 10, Appl
17	23	1.2	897	16	US-10-106-938A-11	Sequence 11, Appl
18	23	1.2	897	16	US-10-106-938A-14	Sequence 14, Appl
19	23	1.2	1281	16	US-10-106-938A-12	Sequence 12, Appl
20	23	1.2	1289	15	US-10-378-393-21	Sequence 21, Appl
21	23	1.2	1939	13	US-10-424-599-1237	Sequence 1237, Ap
22	23	1.2	1987	15	US-10-378-393-17	Sequence 17, Appl
23	23	1.2	4517	16	US-10-093-463-117	Sequence 117, App
24	23	1.2	4586	10	US-09-983-000A-11	GENERAL INFORMATI
25	23	1.2	4586	15	US-10-102-524-1707	Sequence 1707, Ap
26	23	1.2	4620	16	US-10-106-938A-1	Sequence 1, Appli
27	23	1.2	4726	13	US-10-087-192-323	Sequence 323, App
28	23	1.2	144035	13	US-10-087-192-322	Sequence 322, App
29	21	1.1	1179	13	US-10-425-114-4027	Sequence 4027, Ap
30	21	1.1	1186	13	US-10-425-114-4273	Sequence 4273, Ap
31	21	1.1	2196	14	US-10-086-464-7	Sequence 7, Appli
32	21	1.1	2231	14	US-10-086-464-6	Sequence 6, Appli
33	21	1.1	832900	16	US-10-292-798-1393	Sequence 1393, Ap
34	20	1.0	213	9	US-09-867-701-9615	Sequence 9615, Ap
35	20	1.0	260	13	US-10-424-599-58928	Sequence 58928, A
36	20	1.0	432	13	US-09-770-423-122	Sequence 122, App
37	20	1.0	609	9	US-09-800-729-55	Sequence 55, Appl
38	20	1.0	623	13	US-10-425-114-4896	Sequence 4896, Ap
39	20	1.0	655	13	US-10-425-114-11463	Sequence 11463, A
40	20	1.0	727	13	US-10-412-699B-993	Sequence 993, App
41	20	1.0	727	16	US-10-374-780A-492	Sequence 492, App
42	20	1.0	751	13	US-10-425-114-4607	Sequence 4607, Ap
43	20	1.0	767	13	US-10-412-699B-995	Sequence 995, App
44	20	1.0	767	16	US-10-374-780A-494	Sequence 494, App
45	20	1.0	918	13	US-10-424-599-54382	Sequence 54382, A

ALIGNMENTS

RESULT 1

US-10-086-464-1
; Sequence 1, Application US/10086464
; Publication No. US20020199218A1
; GENERAL INFORMATION:
; APPLICANT: GORING, Daphne R. et al.
; TITLE OF INVENTION: PROLINE-RICH EXTENSIN-LIKE RECEPTOR KINASES
; FILE REFERENCE: P 25,762-A USA
; CURRENT APPLICATION NUMBER: US/10/086,464
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 10/069,304
; PRIOR FILING DATE: 2002-02-19
; PRIOR APPLICATION NUMBER: PCT/CA00/00966
; PRIOR FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: US 60/149,466
; PRIOR FILING DATE: 1999-08-19
; PRIOR APPLICATION NUMBER: US 60/159,122
; PRIOR FILING DATE: 1999-10-13
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1944
; TYPE: DNA
; ORGANISM: Brassica napus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1944)
US-10-086-464-1

Query Match 100.0%; Score 1944; DB 14; Length 1944;
Best Local Similarity 100.0%; Pred. NO. 0;
Matches 1944; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 ATGTCTCTCGGCGCGTCTCGGGGACTGTTGGCTTCATCTCCACCATCAACTCCACA 60
|||||
Db 1 ATGTCTCTCGGCGCGTCTCGGGGACTGTTGGCTTCATCTCCACCATCAACTCCACA 60
|||||

QY 61 ACCACCACTCTCTCTCAGCTTCGGCTCTCTCTCCACACACACCTTCTTCTCTCCGCG 120
DB 61 ACCACCACTCTCTCTCAGCTTCGGCTCTCTCTCCACACACACCTTCTTCTCTCCGCG 120
QY 121 CCATCCACTATTCCGACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180
DB 121 CCATCCACTATTCCGACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180
QY 181 CCATCTCCACCAACTCCATCTACGCGGGATCTCCACCTCTCTCTCTCTCTCTCTCTCTCA 240
DB 181 CCATCTCCACCAACTCCATCTACGCGGGATCTCCACCTCTCTCTCTCTCTCTCTCTCTCA 240
QY 241 CCGCTCCAACTACGCGCGGATCTCCACCGGACTCTCTCTCTCTCTCTCTCTCTCTCTCT 300
DB 241 CCGCTCCAACTACGCGCGGATCTCCACCGGACTCTCTCTCTCTCTCTCTCTCTCTCTCT 300
QY 301 CCACCTTCAGTCCCAAGACACCGTCCAACTCTCTCCAGGAGGAGGATCTCTCTGACCT 360
DB 301 CCACCTTCAGTCCCAAGACACCGTCCAACTCTCTCCAGGAGGAGGATCTCTCTGACCT 360
QY 361 CCATCTTCT 420
DB 361 CCATCTTCT 420
QY 421 GGAATCGCCATCGGAGGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 480
DB 421 GGAATCGCCATCGGAGGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 480
QY 481 AAGAAGAAACGAGGAGAGAGAGAGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540
DB 481 AAGAAGAAACGAGGAGAGAGAGAGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540
QY 541 GGTCCCAAGCGGAGGAGCTTACGGTGGACAGCAACAAATGGCGGCAACAAACGCA 600
DB 541 GGTCCCAAGCGGAGGAGCTTACGGTGGACAGCAACAAATGGCGGCAACAAACGCA 600
QY 601 ACACACCGTCAGATCATGTGTGACGTCTACACACACACCTTAAGGCTCCATCTCA 660
DB 601 ACACACCGTCAGATCATGTGTGACGTCTACACACACACCTTAAGGCTCCATCTCA 660
QY 661 CCAGGCAACCTCTCCACCTCCACCGCTTTCATGACGAGCGGGCTCCGAC 720
DB 661 CCAGGCAACCTCTCCACCTCCACCGCTTTCATGACGAGCGGGCTCCGAC 720
QY 721 TACTCGGACCGTCCAGTTCTTCTCTCCACCGTCTCCAGGCTTGTGTAGGCTTCTCCAAA 780
DB 721 TACTCGGACCGTCCAGTTCTTCTCTCCACCGTCTCCAGGCTTGTGTAGGCTTCTCCAAA 780
QY 781 AGCACTTTTCATACGAGGAGCTAGCTAGACCCCAATGGTTCCTCGAGGCGAAGTTG 840
DB 781 AGCACTTTTCATACGAGGAGCTAGCTAGACCCCAATGGTTCCTCGAGGCGAAGTTG 840
QY 841 TTAGGACAAGCGGGTTCGGTTTACGTGCACAAAGTGTGTCCCTAGTGGGAAAGAGTT 900
DB 841 TTAGGACAAGCGGGTTCGGTTTACGTGCACAAAGTGTGTCCCTAGTGGGAAAGAGTT 900
QY 901 GCTGTGAAGCAGTTGAAAGTTGGGAGTGTTCAGGGAGAGGAGTTTCAGGACAGGTT 960
DB 901 GCTGTGAAGCAGTTGAAAGTTGGGAGTGTTCAGGGAGAGGAGTTTCAGGACAGGTT 960
QY 961 GAGATCATCAGCAGATTTCACACAGGATCTGGTGTCTCTGTGGTATTATGATCCGC 1020
DB 961 GAGATCATCAGCAGATTTCACACAGGATCTGGTGTCTCTGTGGTATTATGATCCGC 1020
QY 1021 GGTGCCAAAAGATCTCTGTCTATGATTTGTCTCTTAAACAATCTCGAGCTTCACCTC 1080
DB 1021 GGTGCCAAAAGATCTCTGTCTATGATTTGTCTCTTAAACAATCTCGAGCTTCACCTC 1080
QY 1081 CATGCGGAGGAGCGGCTTACATGGAATGGAGCACCAGATTGAAGATTGCTCTGGATCT 1140
DB 1081 CATGCGGAGGAGCGGCTTACATGGAATGGAGCACCAGATTGAAGATTGCTCTGGATCT 1140
QY 1141 GCTAAAGAGCTTCTTATCTTCATGAAGATTGCAATCTCTAAATCAITTCACCGTGATC 1200

DB 1141 GCTAAAGAGCTTCTTATCTTCATGAAGATTGCAATCTCTAAATCAITTCACCGTGATC 1200
QY 1201 AAGCTTTCAACATATTATGATAGATTTCAGTTTGAAGCTTAAGTTGCTGATTTGGTCTT 1260
DB 1201 AAGCTTTCAACATATTATGATAGATTTCAGTTTGAAGCTTAAGTTGCTGATTTGGTCTT 1260
QY 1261 GCTAAGATTGCTTCTGATACAAACACGCGATGTATCAACGCTGTGATGGGAACCTTTGG 1320
DB 1261 GCTAAGATTGCTTCTGATACAAACACGCGATGTATCAACGCTGTGATGGGAACCTTTGG 1320
QY 1321 TACTTGGCTCCGGAATACGCTGCAAGCGGAAGCTCAAGGAGAGTCTGACGTTTCTCA 1380
DB 1321 TACTTGGCTCCGGAATACGCTGCAAGCGGAAGCTCAAGGAGAGTCTGACGTTTCTCA 1380
QY 1381 TTTGGCGTTGCTTTTGGAGCTCAATCTGGAGCTCGACCCGTTGATGCGCAACAATGTC 1440
DB 1381 TTTGGCGTTGCTTTTGGAGCTCAATCTGGAGCTCGACCCGTTGATGCGCAACAATGTC 1440
QY 1441 TATGTAGATGACAGCTTAGTTGACTGGGACGACCATTTGCTTAACCGAGGATCTGAGCAA 1500
DB 1441 TATGTAGATGACAGCTTAGTTGACTGGGACGACCATTTGCTTAACCGAGGATCTGAGCAA 1500
QY 1501 GGAGACTTTGAGGGTGTAGCTGATGCAAGATGAATATGSGTATGACAGAGAGAGATG 1560
DB 1501 GGAGACTTTGAGGGTGTAGCTGATGCAAGATGAATATGSGTATGACAGAGAGAGATG 1560
QY 1561 GCTCGCATGTTGCTTGTGCTGCGGCTTGTTCGCCATTTCAGCTCGCGCGAGACCTCGC 1620
DB 1561 GCTCGCATGTTGCTTGTGCTGCGGCTTGTTCGCCATTTCAGCTCGCGCGAGACCTCGC 1620
QY 1621 ATGAGCCAGATTGCGTGGCTTGAAGGAATATGATCATTGTCAGATCTTAAACGAAGG 1680
DB 1621 ATGAGCCAGATTGCGTGGCTTGAAGGAATATGATCATTGTCAGATCTTAAACGAAGG 1680
QY 1681 ATGAGCCAGCTCAAGCAATGTATACAGCTCATACGAGGAAGCACCAGATTATGACTCG 1740
DB 1681 ATGAGCCAGCTCAAGCAATGTATACAGCTCATACGAGGAAGCACCAGATTATGACTCG 1740
QY 1741 AGCCAGTACATGAAGACATGAAGAAGTTTAGGAAAATGGCACTTGGAACTCAAGAGTAC 1800
DB 1741 AGCCAGTACATGAAGACATGAAGAAGTTTAGGAAAATGGCACTTGGAACTCAAGAGTAC 1800
QY 1801 AACGCCAGGTGAGTACAGTAATCCGACGAGTACTATGAGACTGTACCCGCTCTGGTTCA 1860
DB 1801 AACGCCAGGTGAGTACAGTAATCCGACGAGTACTATGAGACTGTACCCGCTCTGGTTCA 1860
QY 1861 AGCAGCGAGGCGCAAAACACACGCAATGAGATGGGAAAGATTAAAGAAACCGGTGAG 1920
DB 1861 AGCAGCGAGGCGCAAAACACACGCAATGAGATGGGAAAGATTAAAGAAACCGGTGAG 1920
QY 1921 GGTATAGTGACCTTCTCTTTAA 1944
DB 1921 GGTATAGTGACCTTCTCTTTAA 1944

RESULT 2

US-10-086-464-3
; Sequence 3, Application US/10086464
; Publication No. US20020199218A1
; GENERAL INFORMATION:
; APPLICANT: GORING, Daphne R. et al.
; TITLE OF INVENTION: PROLINE-RICH EXTENSIN-LIKE RECEPTOR KINASES
; FILE REFERENCE: P 25,762-A USA
; CURRENT APPLICATION NUMBER: US/10/086,464
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 10/069,304
; PRIOR FILING DATE: 2002-02-19
; PRIOR APPLICATION NUMBER: PCT/CA00/00966
; PRIOR FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: US 60/149,466
; PRIOR FILING DATE: 1999-08-19
; PRIOR APPLICATION NUMBER: US 60/159,122

; PRIOR FILING DATE: 1999-10-13									
; NUMBER OF SEQ ID NOS: 27									
; SOFTWARE: PatentIn Ver. 2.1									
; SEQ ID NO 3									
; LENGTH: 2189									
; TYPE: DNA									
; ORGANISM: Brassica napus									
; FEATURE:									
; NAME/KEY: CDS									
; LOCATION: (1)..(2189)									
; OTHER INFORMATION:									
US-10-086-464-3									
Query Match 100.0%; Score 1944; DB 14; Length 2189;									
Best Local Similarity 100.0%; Pred. No. 0;									
Matches 1944; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
Qy	1	ATGTCCTCGGCGCGCTCTCCGGGAGCTGGTTCGCTCCATCTCCACCAATCAAACTCCACA	60	841	TTAGGACAAGCGGGTTCCGTTACGTGCACAAAGGTGTGTTCCCTAGTGGGAAAGAGTT	900			
Db	97	ATGTCCTCGGCGCGCTCTCCGGGAGCTGGTTCGCTCCATCTCCACCAATCAAACTCCACA	156	937	TTAGGACAAGCGGGTTCCGTTACGTGCACAAAGGTGTGTTCCCTAGTGGGAAAGAGTT	996			
Qy	61	ACCACCACTCTCTCCAGCTTCCGCTCCTCTCTCCACCAACACCTTCTTCTCCTCCGCGG	120	901	GCTGTGAAGCAGTTTGAAGTTTGGAGTGGTTCAGGAGAGAGGAGTTTTCAGGACAGAGTT	960			
Db	157	ACCACCACTCTCTCCAGCTTCCGCTCCTCTCTCCACCAACACCTTCTTCTCCTCCGCGG	216	997	GCTGTGAAGCAGTTTGAAGTTTGGAGTGGTTCAGGAGAGAGGAGTTTTCAGGACAGAGTT	1056			
Qy	121	CCATCCACTATTTCCGACATCTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	180	961	GAGATCATCAGCAGAGTTTACCACAGGCATCTGCTGTCTCTTGTGTGGTTATTTGCATCGCC	1020			
Db	217	CCATCCACTATTTCCGACATCTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	276	1057	GAGATCATCAGCAGAGTTTACCACAGGCATCTGCTGTCTCTTGTGTGGTTATTTGCATCGCC	1116			
Qy	181	CCATCTCCACCACTCTAGCCGGGATCTCCACCTCTCTCTCTCTCTCTCTCTCTCTCTCA	240	1021	GGTCCCAAAAGATTGCTGTCTATGAGTTTGTCTCTAAACAATCTCGAGCTTCAACCTC	1080			
Db	277	CCATCTCCACCACTCTAGCCGGGATCTCCACCTCTCTCTCTCTCTCTCTCTCTCTCTCA	336	1117	GGTCCCAAAAGATTGCTGTCTATGAGTTTGTCTCTAAACAATCTCGAGCTTCAACCTC	1176			
Qy	241	CCCGCTCCAACTAGCCCGGATCTCCACCGCACTGTACTCTCTCTCTCTCTCTCTCTCT	300	1081	CATGCGAGGAGCGGCTTCAATGGAATGGAGCACCAGATTGAAGATTGCTCTTGGATCT	1140			
Db	337	CCCGCTCCAACTAGCCCGGATCTCCACCGCACTGTACTCTCTCTCTCTCTCTCTCTCT	396	1177	CATGCGAGGAGCGGCTTCAATGGAATGGAGCACCAGATTGAAGATTGCTCTTGGATCT	1236			
Qy	301	CCACTTTCAGTCCAGGACCACTGCTCCACCGCACTGTACTCTCTCTCTCTCTCTCTCT	360	1141	GCTAAAGGACTTCTTATCTTCAAGATTGCAATCTCTAAATCATTCACCGTGATATC	1200			
Db	397	CCACTTTCAGTCCAGGACCACTGCTCCACCGCACTGTACTCTCTCTCTCTCTCTCTCT	456	1237	GCTAAAGGACTTCTTATCTTCAAGATTGCAATCTCTAAATCATTCACCGTGATATC	1296			
Qy	361	CCATCTTCTCCGCGCGCT	420	1201	AAGGCTTCAACATATTTAGATTTCAAGTTTGAAGCTAAGGTTGCTGATTTTGGTCTT	1260			
Db	457	CCATCTTCTCCGCGCGCT	516	1297	AAGGCTTCAACATATTTAGATTTCAAGTTTGAAGCTAAGGTTGCTGATTTTGGTCTT	1356			
Qy	421	GGAATCGGCATCGGAGGAGTGGCTCTGTTGTGATGACTCTGATTTGTTCTCTCTGT	480	1261	GCTAAGATTGCTTCTGATCAAAACAGCATGTATCAACGCTGTGATGGGAACCTTTGGG	1320			
Db	517	GGAATCGGCATCGGAGGAGTGGCTCTGTTGTGATGACTCTGATTTGTTCTCTCTGT	576	1357	GCTAAGATTGCTTCTGATCAAAACAGCATGTATCAACGCTGTGATGGGAACCTTTGGG	1416			
Qy	481	AAGAAGAAACGAG	540	1321	TACTTGTGCTCCGGAATACGCTGCAAGCGGAAAGCTCACGGAAGAGTCTGACGTTTCTCA	1380			
Db	577	AAGAAGAAACGAG	636	1417	TACTTGTGCTCCGGAATACGCTGCAAGCGGAAAGCTCACGGAAGAGTCTGACGTTTCTCA	1476			
Qy	541	GGTCCCAAGCGGAGGAGCTTACGGTGGACAGACAGCAACAAATGGCGCAACAAAGCA	600	1381	TTTGGCGTGTGCTTTTGGAGCTCATTACTGGAGCTCGACCCGTTGATGCAACAATGTC	1440			
Db	637	GGTCCCAAGCGGAGGAGCTTACGGTGGACAGACAGCAACAAATGGCGCAACAAAGCA	696	1477	TTTGGCGTGTGCTTTTGGAGCTCATTACTGGAGCTCGACCCGTTGATGCAACAATGTC	1536			
Qy	601	ACACACCGTCCAGATCATGTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC	660	1441	TATGTAGATGACAGCTTAGTTGACTGGGCACGACCATTTGTTAACCGAGCATCTGAGCAA	1500			
Db	697	ACACACCGTCCAGATCATGTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC	756	1537	TATGTAGATGACAGCTTAGTTGACTGGGCACGACCATTTGCTTAAACGAGCATCTGAGCAA	1596			
Qy	661	CCAGGCAACCTCTCCACCTCCACCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	720	1501	GGAGACTTTTGAGGGTTTGTAGCTGATGCAAGATGAATTAATGGGTATGACAGAGAGATG	1560			
Db	757	CCAGGCAACCTCTCTCCACCTCCACCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	816	1597	GGAGACTTTTGAGGGTTTGTAGCTGATGCAAGATGAATTAATGGGTATGACAGAGAGATG	1656			
Qy	721	TACTCGGACCGTCCAGTTCTTCTCTCCACCGTCTCCAGGGCTTGTGTTAGGGCTTCTC	780	1561	GCTCGCATGGTTGCTGTGCTGGCGCTTGTTCGCCATTCAGCTCGCCGCGAGACCTCGC	1620			
Db	817	TACTCGGACCGTCCAGTTCTTCTCTCCACCGTCTCTCCAGGGCTTGTGTTAGGGCTTCT	876	1657	GCTCGCATGGTTGCTGTGCTGGCGCTTGTTCGCCATTCAGCTCGCCGCGAGACCTCGC	1716			
Qy	781	AGCACTTTTCAATACGAGGAGTAGTAGAGCCCAATGGTTTCTCCGAGGCGAGACTTG	840	1621	ATGAGCCAGATTGTGCGTGGTTAGAGGAAATGTATCATCTGTGATCTTTAAAGAGGG	1680			
Db	877	AGCACTTTTCAATACGAGGAGTAGTAGAGCCCAATGGTTTCTCCGAGGCGAGACTTG	936	1717	ATGAGCCAGATTGTGCGTGGTTAGAGGAAATGTATCATCTGTGATCTTTAAAGAGGG	1776			

```
|||||
Db      2017 GGTATAGTGACCTCTCTTTAA 2040

RESULT 3
US-09-923-876-5096
; Sequence 5096, Application US/09923876
; Patent No. US20020013958A1
; GENERAL INFORMATION:
; APPLICANT: Kamigaki, Laura Y. (Ito)
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN SEEDLING
; FILE REFERENCE: PL-0012-1 CON
; CURRENT APPLICATION NUMBER: US/09/923,876
; CURRENT FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: 09/298,329
; PRIOR FILING DATE: 1999-04-21
; PRIOR APPLICATION NUMBER: 60/085,331
; PRIOR FILING DATE: 1998-05-05
; NUMBER OF SEQ ID NOS: 6332
; SOFTWARE: PERL Program
; SEQ ID NO 5096
; LENGTH: 268
; TYPE: DNA
; ORGANISM: Zea mays
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020013958A1 700456354H1
; NAME/KEY: unsure
; LOCATION: 63
; OTHER INFORMATION: a, t, c, g, or other
US-09-923-876-5096

Query Match      1.3%; Score 26; DB 9; Length 268;
Best Local Similarity 100.0%; Pred. No. 0.0045;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1231 TTGAAGCTAAGGTTGCTGATTTGG 1256
|||||
Db      203 TTGAAGCTAAGGTTGCTGATTTGG 228

RESULT 4
US-09-923-876-5096
; Sequence 5096, Application US/09923876
; Publication No. US20030237110A9
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Kamigaki, Laura Y. (Ito)
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN SEEDLING
; FILE REFERENCE: PL-0012-1 CON
; CURRENT APPLICATION NUMBER: US/09/923,876
; CURRENT FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: 09/298,329
; PRIOR FILING DATE: 1999-04-21
; PRIOR APPLICATION NUMBER: 60/085,331
; PRIOR FILING DATE: 1998-05-05
; NUMBER OF SEQ ID NOS: 6332
; SOFTWARE: PERL Program
; SEQ ID NO 5096
; LENGTH: 268
; TYPE: DNA
; ORGANISM: Zea mays
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20030237110A9 700456354H1
; NAME/KEY: unsure
; LOCATION: 63
; OTHER INFORMATION: a, t, c, g, or other
US-09-923-876-5096

Query Match      1.3%; Score 26; DB 9; Length 268;
Best Local Similarity 100.0%; Pred. No. 0.0045;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1231 TTGAAGCTAAGGTTGCTGATTTGG 1256
|||||
Db      203 TTGAAGCTAAGGTTGCTGATTTGG 228

RESULT 5
US-10-425-114-1922
; Sequence 1922, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 1922
; LENGTH: 951
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700171392_FLI
US-10-425-114-1922

Query Match      1.3%; Score 26; DB 13; Length 951;
Best Local Similarity 100.0%; Pred. No. 0.0046;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1231 TTGAAGCTAAGGTTGCTGATTTGG 1256
|||||
Db      16 TTGAAGCTAAGGTTGCTGATTTGG 41

RESULT 6
US-10-425-114-13131
; Sequence 13131, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 13131
; LENGTH: 1011
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: JC-GMFL02220096D10_FLI
US-10-425-114-13131

Query Match      1.3%; Score 26; DB 13; Length 1011;
Best Local Similarity 100.0%; Pred. No. 0.0046;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1360 GAGAAGCTGACGTTTCTCATTGG 1385
```



```
Db      246 GAGAAGTCTGACGTTTCTCATTTGG 271
|||||
RESULT 7
US-10-425-114-36420
; Sequence 36420, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 36420
; LENGTH: 1134
; TYPE: DNA
; ORGANISM: Zea mays subsp. mexicana
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMROTBOSINTE097B02_FLI
US-10-425-114-36420

Query Match      1.3%; Score 26; DB 13; Length 1134;
Best Local Similarity 100.0%; Pred. No. 0.0047;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1231 TTGAAGCTAAGGTTGCTGATTTGG 1256
|||||
Db      260 TTGAAGCTAAGGTTGCTGATTTGG 285

RESULT 8
US-10-425-114-6300
; Sequence 6300, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 6300
; LENGTH: 1500
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700560157_FLI
US-10-425-114-6300

Query Match      1.3%; Score 26; DB 13; Length 1500;
Best Local Similarity 100.0%; Pred. No. 0.0047;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1360 GAGAAGTCTGACGTTTCTCATTTGG 1385
|||||
Db      779 GAGAAGTCTGACGTTTCTCATTTGG 804

RESULT 9
US-10-425-114-599-47702
; Sequence 47702, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 47702
; LENGTH: 1552
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_143080C.1
US-10-424-599-47702

Query Match      1.3%; Score 26; DB 13; Length 1552;
Best Local Similarity 100.0%; Pred. No. 0.0047;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1360 GAGAAGTCTGACGTTTCTCATTTGG 1385
|||||
Db      779 GAGAAGTCTGACGTTTCTCATTTGG 804

RESULT 10
US-10-425-114-2326
; Sequence 2326, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 2326
; LENGTH: 1724
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700209610_FLI
US-10-425-114-2326

Query Match      1.3%; Score 26; DB 13; Length 1724;
Best Local Similarity 100.0%; Pred. No. 0.0047;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1231 TTGAAGCTAAGGTTGCTGATTTGG 1256
|||||
Db      832 TTGAAGCTAAGGTTGCTGATTTGG 857

RESULT 11
US-10-424-599-122411
; Sequence 122411, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
```

```
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 122411
; LENGTH: 1101
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MFT3847_81542C.1
US-10-424-599-122411

Query Match 1.3%; Score 25; DB 13; Length 1101;
Best Local Similarity 100.0%; Pred. No. 0.017;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1178 CTAATAATCATTCACCGTGATATCAA 1202
|||||
Db 69 CTAATAATCATTCACCGTGATATCAA 93

RESULT 12
US-10-368-046-1
; Sequence 1, Application US/10368046
; Publication No. US20040063188A1
; GENERAL INFORMATION:
; APPLICANT: Robin A. Robinson
; APPLICANT: Vittoria Cloce
; TITLE OF INVENTION: Method for Isolation and Purification of
; TITLE OF INVENTION: Expressed Gene Products In Vitro
; FILE REFERENCE: 44149-3US1
; CURRENT APPLICATION NUMBER: US/10/368,046
; CURRENT FILING DATE: 2003-02-15
; PRIOR APPLICATION NUMBER: US 60/356,119
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/356,161
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/356,118
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/356,133
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/356,157
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/356,156
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/356,123
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/356,113
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/356,154
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/356,135
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1515
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HPV-16 L1 codon optimized sequence
US-10-368-046-1

Query Match 1.2%; Score 24; DB 13; Length 1515;
Best Local Similarity 100.0%; Pred. No. 0.059;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 669 ACCTCTCCACCTCCACCGCC 692
|||||
Db 1471 ACCTCTCCACCTCCACCGCC 1494

RESULT 14
US-10-029-386-26494
; Sequence 26494, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 26494
; LENGTH: 180
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
```

```
Db 1471 ACCTCTCCACCTCCACCGCC 1494

RESULT 13
US-10-367-095-1
; Sequence 1, Application US/10367095
; Publication No. US20030228696A1
; GENERAL INFORMATION:
; APPLICANT: Robin A. Robinson
; TITLE OF INVENTION: No. US20030228696A1el Insect Cell Line
; FILE REFERENCE: 44149-1US1
; CURRENT APPLICATION NUMBER: US/10/367,095
; CURRENT FILING DATE: 2003-02-14
; PRIOR APPLICATION NUMBER: US 60/356,119
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/356,161
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/356,118
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/356,133
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/356,157
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/356,156
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/356,123
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/356,113
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/356,154
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/356,135
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1515
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HPV-16 L1 codon optimized sequence
US-10-367-095-1

Query Match 1.2%; Score 24; DB 16; Length 1515;
Best Local Similarity 100.0%; Pred. No. 0.059;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 669 ACCTCTCCACCTCCACCGCC 692
|||||
Db 1471 ACCTCTCCACCTCCACCGCC 1494

RESULT 14
US-10-029-386-26494
; Sequence 26494, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 26494
; LENGTH: 180
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
```

OTHER INFORMATION: MAP TO CHR7.1
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.58
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.59
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.45
OTHER INFORMATION: SWISSPROT HIT: P16056, EVALUE 1.00e-27
OTHER INFORMATION: NT HIT: g1455746, EVALUE 8.00e-90
OTHER INFORMATION: EST_HUMAN HIT: AW418529.1, EVALUE 8.00e-97
US-10-029-386-26494

Query Match 1.2%; Score 23; DB 15; Length 180;
Best Local Similarity 100.0%; Pred. No. 0.2;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1240 AAGTTGCTGATTTGGTCTTGC 1262
|||||
Db 29 AAGTTGCTGATTTGGTCTTGC 51

RESULT 15
US-10-029-386-12794
Sequence 12794, Application US/10029386
Publication No. US20030194704A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
FILE REFERENCE: AEMICA-X-2
CURRENT APPLICATION NUMBER: US/10/029,386
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 34288
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 12794
LENGTH: 579
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO CHR7.1
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.58
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.59
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.45
OTHER INFORMATION: NT HIT: U96996.1, EVALUE 1.00e-107
OTHER INFORMATION: SWISSPROT HIT: P16056, EVALUE 6.00e-27
OTHER INFORMATION: EST_HUMAN HIT: AW418529.1, EVALUE 0.00e+00
US-10-029-386-12794

Query Match 1.2%; Score 23; DB 15; Length 579;
Best Local Similarity 100.0%; Pred. No. 0.21;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1240 AAGTTGCTGATTTGGTCTTGC 1262
|||||
Db 221 AAGTTGCTGATTTGGTCTTGC 243

Search completed: May 12, 2004, 22:42:51
Job time : 843 secs

This page Blank (uspto)